Lytic Polysaccharide Monooxygenases - Studies of Fungal Secretomes and Enzyme Properties

Nekiunaite, Laura

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

LYTIC POLYSACCHARIDE MONOOXYGENASES – STUDIES OF FUNGAL SECRETOMES AND ENZYME PROPERTIES

PhD thesis

Laura Nekiunaite
March 2016

DTU Systems Biology
Department of Systems Biology

Preface

The current PhD thesis presents the results of my work carried out in the Enzyme and Protein Chemistry (EPC) group, Department of Systems Biology, Technical University of Denmark from December 2012 to March 2016 under supervision of Associate Professor Maher Abou Hachem, Associate Professor Gustav Vaaje-Kolstad (NMBU, Norway) and Professor Birte Svensson. The project was supported by the Novo Nordisk foundation with a "Biotechnology-based Synthesis and Production" grant (NNF12OC0000769).

The work has resulted in the following manuscripts:

Laura Nekiunaite, Magnus Ø. Arntzen, Birte Svensson, Gustav Vaaje-Kolstad and Maher Abou Hachem. Lytic polysaccharide monooxygenases and other oxidative enzymes are abundantly secreted by *Aspergillus nidulans* grown on different starches. *In preparation*.

Laura Nekiunaite, Trine Isaksen, Gustav Vaaje-Kolstad and Maher Abou Hachem. A starch specific fungal lytic polysaccharide monooxygenase binds to starch with similar affinity as amylolytic hydrolases. *In preparation*.

Laura Nekiunaite, Dejan Petrovic, Bjørge Westereng, Gustav Vaaje-Kolstad, Maher Abou Hachem, Aniko Varnai, Vincent G.H. Eijsink. The *Fusarium graminearum* lytic polysaccharide monooxygenase *Fg*LPMO9A cleaves xyloglucan independent on the backbone substitution pattern. *In preparation*.

Laura Nekiunaite, Kgs. Lyngby, March 2016

Acknowledgements

This thesis is not a result of work done by me alone, but has been accomplished with the help and support of a number of people whom contributed in many ways.

First, I would like to thank Maher Abou Hachem for supervising my PhD for the past 3 years and for giving me the opportunity to study this exciting field of research. Likewise, I would like to thank my co-supervisors, Gustav Vaaje-Kolstad, for a welcoming stay in Norway and for his positive attitude and encouragements, and Birte Svensson, for sharing her knowledge and expert advice. I would like to express my gratefulness to Bjørge Westereng for introducing me to Maldi, explaining and discussing all the things that were new to me. I would also like to thank Jakob B. Nielsen for introducing me to world of *Aspergillus*.

Many people at DTU made my PhD journey much more pleasant. I am grateful to all former and members of EPC for their help. My special thanks to Susan and Eva for sharing the office with me, all the discussions that were not all scientific, for their friendship and time spend together. I also would like to acknowledge previous PhD student Alexander for always being willing to help, to discuss science, and for all our nice time running and biking together.

I would like to thank all the people from PEP group at NMBU in Norway. Everybody helped making my two long stays away from home much easier. First, thanks to Vincent G.H. Eijsink for welcoming me into his group. Thanks to Trine for all the cozy chats. Thanks to Magnus for all the help with secretomics and for keeping me busy with all data files, Aniko and Dejan, for all the help with *Fusarium* enzyme, and Zarah for always being so helpful expert in the lab. Thanks to Live, Kasia, Tina, Ben, Leszek, Jeremy, Daw and many others (it's such a big group) for all the entertainment apart from work.

Last, but by no means least, I would like to thank my parents, Johannes and his family, Ruta and other friends for all the support and believing in me during all this time. Thanks for listening to me through all my ups and downs, and for listening to those strange words like monooxygenases, which you still claim not to know what they mean.

Summary

Efficient degradation of plant biomass by enzymes is an important step towards a more environmentally friendly and sustainable bioeconomy. However, the complexity and recalcitrant nature of the substrates limit enzyme performance on plant biomass and current enzyme cocktails are not efficient enough to degrade it. The recently discovered lytic polysaccharide monooxygenases (LPMOs) are crucial enzymes employed in biomass breakdown in nature owing to their ability to boost activity of other biomass degrading hydrolases.

Filamentous fungi are known to be significant players in plant biomass conversion as they produce a wide diversity of degrading enzymes. In the first part of this PhD thesis, the secretomes of the well-known fungus *Aspergillus nidulans* grown on cereal and legume starches were analyzed. Secretomics is a powerful tool to unravel secretion patterns of fungi and their response to different substrates at the protein level. It could help to design better enzyme cocktails that increase efficiency of biomass degradation. The secretomes of *A. nidulans* revealed differences in growth and secretion of enzymes, depending on the type and properties of starches. A common characteristic of the fungus secretomes on different starches was that the LPMOs, shown to be active on starch, were highly abundant, together with other oxidative enzymes suggesting an important role for these enzymes in fungal starch degradation. The presence of binding sites for AmyR, a transcriptional regulator for starch degradation, were also identified upstream the LPMO genes, providing evidence for a co-regulatory mechanism of LPMOs and amylolytic hydrolases.

The second part of the PhD thesis is focused on understanding the binding properties of LPMOs to starch and starch mimic substrate. It was shown that LPMOs possessing starch binding domain (CBM20) has similar binding affinities as the amylolytic hydrolases containing CBM20. The binding of the LPMO to starch is presumably mediated by the CBM20.

In the last part of the PhD thesis, the action of a novel LPMO from the plant pathogenic fungus *Fusarium graminearum* on cellulose and xyloglucan is demonstrated. Previous studies have shown that some fungal LPMOs are capable to degrade xyloglucan but only by unsubstituted glucose unit in the backbone. This study for the first time showed that LPMO from *F. graminearum* is able to cleave xyloglucan backbone randomly, including two substituted glucose units. Moreover, a question about a connection between LPMOs and fungi pathogenesis was raised.

The novel insight from work done during the PhD study promotes our understanding of unexplored aspects of fungal LPMOs and inspires further research in this realm of glycoscience.

Dansk resume

Effektiv enzymatisk nedbrydning af plantebiomasse er et vigtigt skridt på vejen mod en mere miljøvenlig og bæredygtig bioøkonomi. Eksisterende enzymblandinger er imidlertid ikke tilstrækkeligt velegnede til at nedbryde denne type af meget komplekse og modstandsdygtige substrater. Man har imidlertid for nyligt opdaget lytic polysaccharide monooxygenaser (LPMOs), som har vist sig at spille en vigtig rolle i naturen takket være deres evne til at forøge den reelle aktivitet af kulhydrathydrolaser, der traditionelt er kendt for at nedbryde plantebiomasse.

Filamentøse svampe tilskrives en vigtig rolle i omdannelse af plantebiomasse, da de producerer en bred vifte af nedbrydende enzymer. I første del af denne ph.d. afhandling analyseres sekretomer fra *Aspergillus nidulans*, der er en velkendt filamentøs svamp, som dyrkes på stivelse fra korn og bælgfrugter. Sekretomanalyse er et værktøj der på proteinniveau kan kortlægge enzymproduktionsmønstre for svampe og ved hjælp af massespektrometri identificere enzymer der virker på de undersøgte substrater. På grundlag af resultater fra sekretomsanalyse bliver det muligt at foreslå bedre enzymblandinger til nedbrydning af plantebiomasse. Sekretomer fra *A. nidulans* viste enzymprofiler afhængighed af dyrkningsmediets stivelsessubstrat. Fælles for svampesekretomer opnået med forskellige stivelsessubstrater var indhold af et større antal LPMOer, som er kendt for at være aktive over for stivelse, samt indhold af andre oxidative enzymer. Dette peger på at disse enzymer spiller en vigtig rolle i stivelsesnedbrydning af filamentøse svampe. Bindings sites for AmyR, der er transcriptionel regulator for nedbrydning af stivelse, blev identificeret i generne for LPMO, hvilket underbyggede, at der findes en helt eller delvis fælles regulatorisk mekanisme for LPMOer og stivelsesnedbrydende hydrolaser.

I anden del af ph.d. afhandlingen fokuseres på forståelsen af LPMOers binding til stivelse og stivelses-lignende kulhydrater. Det er vist, at LPMOer, som indeholder et stivelsesbindende domæne fra *"carbohydrate binding module"* familie 20 (CBM20) har bindingsegenskaber, som svarer til amylolytiske hydrolaser indeholdende CBM20. Formodentlig medieres LPMO's binding til stivelse af CBM20.

I den sidste del af ph.d. projektet beskrives effekten af nye LPMOer fra en plantepatogen *Fusarium graminearum* overfor cellulose og xyloglukan. Tidligere undersøgelser har vist, at nogle LPMOer fra svampe kan bruges til at nedbryde xyloglukan, men at spaltningen kun sker ved ikke substituerede glukoseenheder. Jeg har i mit arbejde opdaget at LPMOer fra *F. graminearum* kan kløve mellem tilfældige glukoserester i xyloglukan også selvom de er substituerede med andre

sukkerenheder. Dette resultat rejser et spørgsmål om sammenhængen mellem LPMOer og patogene svampe.

Undersøgelser udført i forbindelse med dette ph.d. projekt har øget kendskabet til uudforskede aspekter ved stivelsesnedbrydende LPMOer fra filamentøse svampe og inspirerer til fremtidig forskning inden for dette område af glyko-videnskab.

Abbreviations

AA	Auxiliary activities
CAZyme	Carbohydrate-active enzyme
CBM	Carbohydrate-binding module
CDH	Cellobiose dehydrogenase
CE	Carbohydrate esterase
CM	Catalytic module
DP	Degree of polymerization
EPR	Electron paramagnetic resonance
GH	Glycoside hydrolase
GT	Glycoside transferase
LPMO	Lytic polysaccharide monooxygenase
MS	Mass spectrometry
ORF	Open reading frame
PL	Polysaccharide lyase
XG	Xyloglucan

Contents

CHAPTER 1. INTRODUCTION

1 Bioeconomy and biomass

The wealth and development of mankind demands an increasing energy production, which nowadays mainly relies on fossil fuel sources. Issues like CO_2 emission and global climate change, in combination with decreasing fossil fuel reserves, have been driving the search for alternative and renewable energy sources [1]. The European Commission has presented a long-term framework for developing a competitive low carbon economy for the next few decades [2,3]. A part of the sustainable economy refers to the sustainable production and conversion of biomass into a range of food, health, materials and other industrial products and not the least, energy. Moreover, the transition towards a modern bioeconomy faces various challenges such as the sustainability of biomass raw material, efficiency in biomass use and economy of biomass deployment [2,4].

Biomass is defined as organic matter derived from living, or recently living organisms (Table 1). This mostly refers to plants and plant-derived material that can be used as a feedstock for biofuels or biopower production, and bioproducts that would otherwise be made from fossil fuels. Based on the source of biomass, biofuels are classified broadly into two major categories: first- and second-generation biofuels. First-generation biofuel is derived from sources such as sugarcane, grains and vegetable oils, which offers some CO_2 emission reduction benefits and can help to improve domestic energy security. However, there are concerns regarding utilizing food-based resources for its environmental impact and the ethical and economic issues related to the impact of this on food and feed supplies [1,5]. On the other hand, second-generation biofuels use non-food-based biomass resources that mainly consist of lignocellulosic material, e.g. trees and forestry co-products or agricultural residues like straw or bagasse. Even though it is a more favored substitute, technological issues and uncompetitive economy constitute major hinders in the transition towards this type of fuel [1,6]. There are also two new categories (third- and fourth-generation) but they are still at a very early stage of development [7].

Table 1. Typical biomass derived feedstocks

Biomass feedstocks[a]	**Examples**[a]
Non-woody	Sugar cane, corn, wheat, sugar beet, algae, etc.
Woody	Logging residues, forest thinnings, mill residues, etc.
Animal byproducts	Tallow, fish oil, manure, slaughterhouse water, etc.
Agricultural residues	Wheat straw, bagasse, corn stover, rice husk, etc.
Wastes	Manufacturing waste, municipal waste, landfill gas, sewage sludge, waste cooking oil, etc.

[a] Adapted from [2,3,5,8–10].

The future prospect of using biomass in various processes poses both technical and economic challenges. One of the main issues is related to the recalcitrance of lignocellulosic feedstocks. Moreover, the performance of the enzymes used for depolymerization of biomass is currently not efficient and thus affects conversion costs [6,11]. It is clear that technological advances must be realized to make biofuels sustainable and cost effective. Efforts to improve the efficiency of biomass pretreatment and enzymatic hydrolysis methods are in focus. Therefore, discovery of new enzymes and new enzyme cocktails that could improve enzymatic deconstruction of feedstocks is likely to promote the feasibility of biomass conversion processes. In this chapter, the structures of major plant biomass polymers are presented together with an overview of the remarkable capabilities evolved by fungi to degrade these polymers, and the fungal enzymes that confer this breakdown.

2 Plant biomass major polysaccharides

Plant biomass, which is one of the most abundant and renewable biological resources worldwide, is seen as a promising feedstocks for fuels and raw materials. It is produced from photosynthesis, where atmospheric CO_2 is converted to carbohydrates using solar energy. It is estimated that primary biomass synthesis by plants is about 140 billion tons annually [12].

The main plant biomass components are polysaccharides and lignin, which is a heterogeneous aromatic polymer. Plant polysaccharides are either serving a structural role in the plant cell wall or as storage reservoirs. The major polysaccharides serving as energy reserves are starch followed by inulin and mannans that are found in seeds [13]. There are two major types of cell walls in plants – the primary cell wall (a layer that is produced by growing cells) and the secondary cell wall (a layer that is formed inside the primary cell wall after termination of cell growth). Plant cell walls are composed primarily of cross-linked polysaccharides, cellulose and hemicellulose; the primary cell wall also contain pectins as well as structural proteins, whereas the thicker secondary wall contains little protein or pectin, but normally contains lignin [14–16].

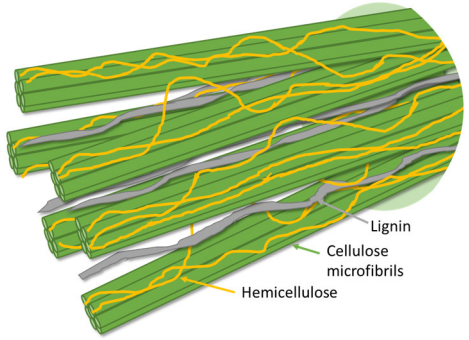


Figure 1. Schematic model for the arrangement of load-bearing cellulose microfibrils (green cylinders) coated with hemicellulose chains (yellow) and lignin (grey) in the plant cell walls. Adapted from [17].

Hemicelluloses coat the cellulose microfibrils (Figure 1) and may be covalently bound to lignin. These three components together are known as lignocellulose. This close association may prevent aggregation of cellulose microfibrils and enable the interactions between cellulose microfibrils and other cell wall polymers, thereby strengthening the mechanical rigidity of the plant cell wall [18,19]. Depending on plant species and cell type, the dry weight of a cell wall typically consists of about 35–50% cellulose, 20–35% hemicellulose, and 10–25% lignin [20].

2.1 Cellulose

Considering plant biomass, the main component is the cellulose (Figure 2), the most abundant biopolymer on earth. Cellulose is mainly associated with plants as they represent the major sources of polysaccharide. However, some other organisms also produce cellulose, including a variety of bacteria, fungi, algae, and even some animals (tunicates) [21,22].

Cellulose is the dominant structural polysaccharide of the plant cell wall. Chemically, cellulose is a linear β-1,4-linked insoluble homopolymer of D-glucose, with the disaccharide cellobiose as its repeating unit (Figure 2). The degree of polymerization (DP) of the β-1,4 D-glucan polymer chains of cellulose varies from several hundreds to tens of thousands depending on the origin, and a DP above 8 renders the glucan chain insoluble [21]. Cellulose polymers are arranged into microfibrils that aggregate to form a complex assembly of cellulose fibers [22,23].

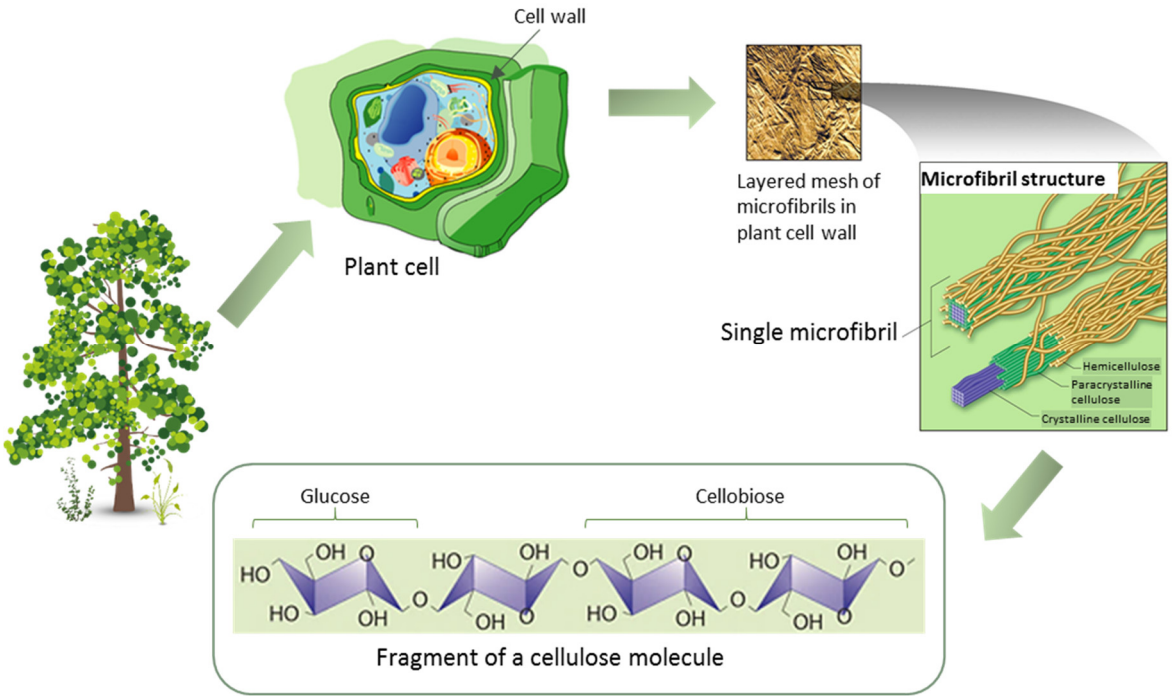


Figure 2. Schematic representation of different structural levels of cellulose: from cellulose fibers in the plant cell wall to atomic level. Adapted from [24].

Native cellulose consists of crystalline arrays interrupted by amorphous regions [25]. The amorphous regions in cellulose likely arise from chain dislocations on segments where fibrils are distorted along their length due to internal strain [26]. In the crystalline cellulose, the chains are tightly packed together and stabilized by a strong intra- and intermolecular hydrogen-bonding

networks. Cellulose can pack into multiple crystalline forms, or polymorphs, which differ in symmetry and chain geometry. Native cellulose is cellulose I, that occurs in two crystalline forms: the more rare Iα found in bacteria and algae, and Iβ that is dominant in higher plants [22,27]. Both of these crystal forms share the same conformation, with a parallel-oriented chain arrangement but differ in hydrogen bonding patterns and the interlayer chain stacking arrangement.

2.2 Hemicellulose

Hemicellulose is one of the major parts of plant biomass and an important component in food and feed. Hemicelluloses are very complex heteropolysaccharides constituted of pentoses (xylose, arabinose), hexoses (mannose, glucose, galactose) and sugar acids (glucuronic acid). The monosaccharide moieties in the backbone chains of hemicellulose are commonly joint with β-1,4-glycosidic linkages (except for mixed linkage β-glucans) and all have the same equatorial configuration at C1 and C4. Hemicelluloses include xyloglucans, heteroxylans, heteromannans and mixed-linkage β-glucans [28]. The composition of hemicellulose is highly variable between different plants and plant cells. These polysaccharides are different from each other structurally, some of them are linear, whereas other ones are more complex, containing branches and chemical modifications of certain sugars. The main hemicellulose structures are presented in Figure 3.

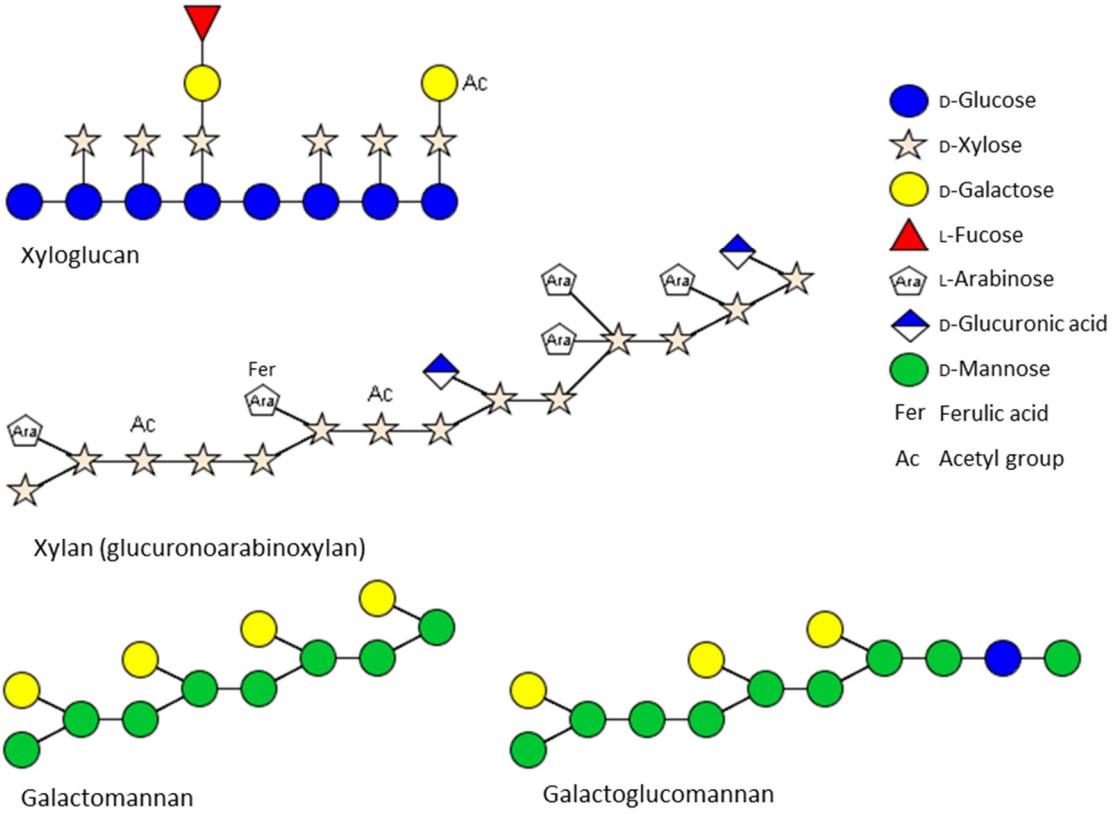


Figure 3. Schematic representation of common hemicellulose structures found in the plant cell walls: xyloglucan, xylan, galactomannan and galactoglucomannan. The structure of the hemicelluloses varies greatly in different plant species and tissue types. Adapted from [28].

Xyloglucan (XG) is an important hemicellulose in a large proportion of the plant kingdom as both structural and storage polysaccharide [29,30]. The localization of XG is the middle lamella, primary walls, and gelatinous wall layer of higher plants, while the abundance of this polymer varies depending on the source: from 25% (dry weight) in the primary cell walls of flowering plants to less than 2% in cereals and other grasses [31–33]. The backbone of XG is β-1,4-glucan that is heavily substituted with pendant α-1,6-xylosyl units, and this structural core is shared in all plants (Figure 3). Additional substituents are frequently observed depending on plant family, tissue, cell type and developmental stage, and the branching pattern has both functional and taxonomic significance. In many species, the backbone has a pattern of three substituted glucose units followed by an unsubstituted glucose residue [34]. The xylosyl residue may be linked with a D-galactosyl residue which may be additionally linked to an L-fucosyl residue (referred to as side chain F: α-L-Fuc*p*-1,2-β-D-Gal*p*-1,2-α-D-Xyl*p*-1,6-β-D-Glc*p*), and also other side chain patterns have been found [18,30]. In addition, some XG molecules can be acetylated at the galactosyl moiety [18]. XG and cellulose

microfibrils have been proposed to form a network that has a major regulatory and load-bearing functions in the plant primary cell walls [34]. XG coats most of the available cellulose microfibril surfaces by intercalating within microfibrils and tethering them (Figure 1) [28,34].

Xylans are the main hemicellulose found in the primary and secondary cell walls of grasses, as well as in the secondary cell walls of hardwood trees [35]. Xylan is composed of β-1,4-linked xylose backbone, which can be substituted with arabinofuranose, glucuronic acid and ferulic acid moieties (Figure 3), depending on the species [28]. Most xylans are also acetylated [35]. Xylan found in monocots is mainly glucuronoarabinoxylan, and in dicots – glucuronoxylans [28].

Another dominant hemicellulose in grasses, taking up to 30% (dry weight) of their primary cell wall, is mixed-linkage β-glucans [33,36]. Mixed-linkage glucans are unbranched polymers mainly composed of β-1,4-linked cellotriosyl and cellotetrasyl units connected by β-1,3 linkages [35].

Several types of polysaccharides contain β-1,4-linked mannan in their backbone: mannan, galactomannan, glucomannans and galactoglucomannan. Mannans and galactomanns consist of a mannose backbone, whereas glucomannan and galactoglucomannans consist of alternating mannose and glucose units in their backbone in a nonrepeating arrangement (Figure 3) [28]. Mannosyl residues in glucomannan and galactoglucomannans can be substituted with galactosyl residues by α-1,6 linkages. [28,37].

2.3 Starch

Starch, the second most abundant biomass material in nature after cellulose, is a natural, renewable polymer produced by all plant species as a source of stored energy either in transitory form in leaves or for long term storage in nonphotosynthetic organs such as roots, tubers and seeds [38,39]. Starch is biosynthesized as granules within the cell in specialized compartments, the plastids [40]. Besides grains, such as wheat and corn, and tubers, such as cassava, abundant starch-based feedstocks, e.g. wheat bran, potato peels, rice husks, wasted crop and brewery spent grains, also could be used as a low-cost residual biomass for the production of bioethanol [41].

Starch is a semi-crystalline, insoluble homo-glucose polymer. It is composed of two main components: amylopectin and amylose (Figure 4). Amylopectin is the major component of the starch granule, building up to 70–80% of the granules in most plants [40]. It is a highly branched molecule consisting of α-1,4-linked linear chains and 5–6% side chains linked by α-1,6 branch points [38]. Amylose is an essentially linear polymer consisting of α-1,4-linked glucose units with very few α-1,6 branches. While normally amylose content in starch is around 20–30%, in different types of genetically altered starches it can be between 0 and 80 %. The degree of polymerization varies

among and within plant species. In wheat, the DP of amylose was reported to be from 830 to 1570 [42], while the DP of amylopectin is higher than for corresponding values for amyloses, varying up to 26500 [40].

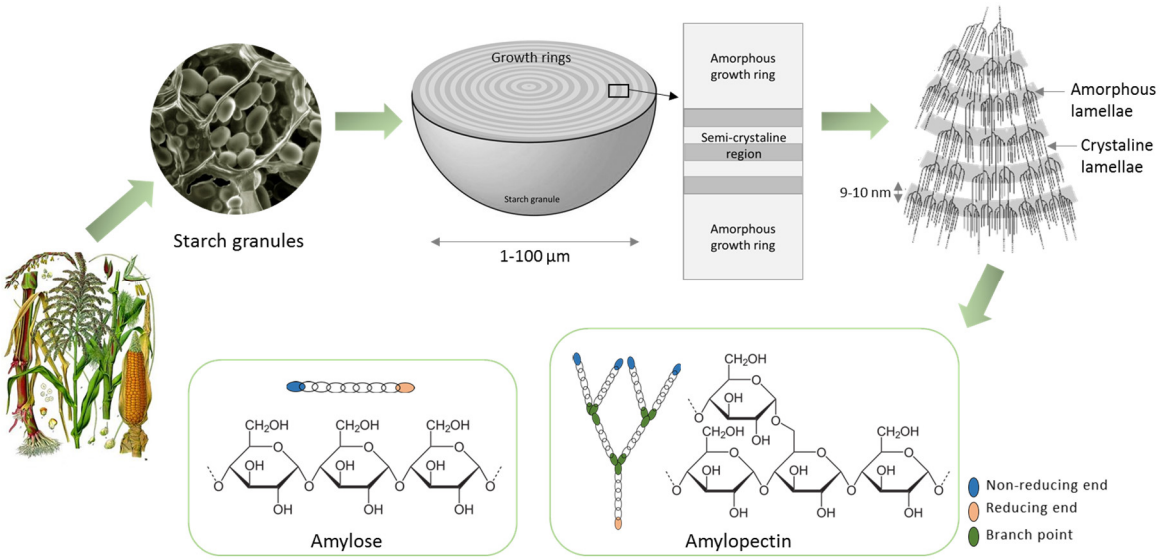


Figure 4. Schematic representation of different structural levels of the starch granule: from granular to atomic level. Adapted from [38,43].

Starch is deposited as semi-crystalline granules that vary from 15 to 45% in their crystallinity [40,44]. Crystallinity is associated with the amylopectin component and is critical concerning starch granule architecture and physicochemical characteristics like the susceptibility to enzymes and the insolubility in cold water [45]. The inner architecture of most starch granules is characterized by growth rings, alternating amorphous and semi-crystalline shells which are 100–400 nm thick (Figure 4) [46]. Within the semi-crystalline shells there are crystalline lamellae, formed by clusters of double helical adjacent glucan chains in amylopectin, and amorphous lamellae, formed by the regions where branch points of amylopectin occur (Figure 4) [40]. Amylose is distributed in the amorphous regions. The arrangement of lamellar is periodically repeated every 9 to 10 nm (Figure 4) [39].

Table 2. Characteristics of starch granules and amylose content in different plant sources

STARCH	Type	Amylose content[a], %	Diameter, µm	Shape
Barley	Cereal	23	A[b]: 15–25 B[c]: 2–3	Lenticular (A) Spherical (B)
Wheat	Cereal	25	A[b]: 20–35 B[c]: 2–3	Lenticular (A) Perfect spheres (B)
Maize	Cereal	25	2–30	Polyhedral/spherical
High amylose maize	Cereal	70	5–25	Elongated irregular filament
Rice	Cereal	20	3–8	Polyhedral
Potato	Tuber	20	15–100	Lenticular
Pea	Legume	35	10–45	Oval

Adapted from [38,47].
[a] average amylose content; [b] A granule – large granule; [c] B granule – small granule.

The organization of granules is very dependent on the botanical origin and therefore the granules differ in their size, morphology, composition, and degrees of crystallinity (Table 2). There is a distribution of granule size between each plant species and even within each plant: diameter ranging from 2 to 100 µm. Varying granule shape (spherical, oval, lenticular) demonstrates that the biosynthetic pathway together with the physical constraints during cellular starch deposition can shape granules. Most likely, the crystallinity of granules depends on both genetic control and climate conditions during the tuber growth and grain filling periods of the plants [38].

3 Bioprocessing of plant biomass

The world is in a transition from fossil-based towards a bio-based economy. Therefore, more efficient and sustainable ways of utilizing renewable resources need to be established. The biorefinery is a promising concept that integrates abundant raw material conversion processes to produce a mixture of products like biofuels, bioproducts, chemicals and direct energy (Figure 5) [8,48,49]. For example, the Borregaard biorefinery in Norway is one of the most advanced and successful biorefineries in the world producing advanced and environmentally friendly bioproducts (like specialty cellulose, vanillin, bioethanol and lignin based products) that can replace oil-based products by using sustainable wood-based raw materials.

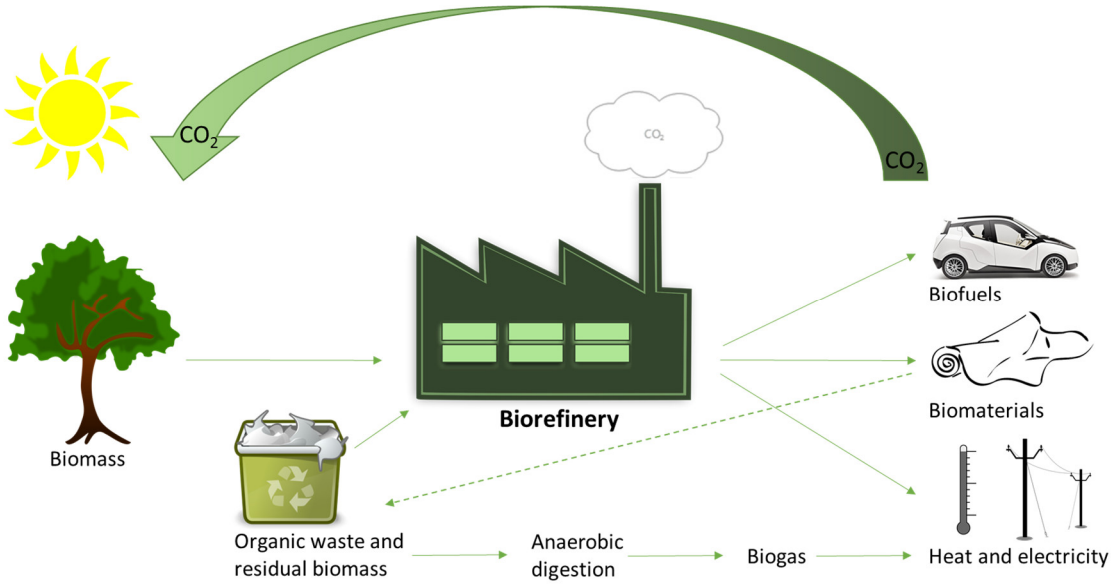


Figure 5. The biorefinery concept based on plant biomass feedstocks. Adapted from [48].

At present, harsh chemical treatments and high temperatures are still used in industry to deconstruct plant derived polymers, specifically plant cell walls [50]. One of the most promising approaches for biomass deconstruction is biochemical and relies on the use of enzymes [50]. The enzymatic deconstruction of biomass is generally considered as sustainable and therefore enzymes are often preferred over inorganic compounds. Enzymes are very attractive tools for modern, green biotechnological applications due to their high selectivity, milder reaction conditions e.g. temperature and pH, and possibilities for recirculation. However, comprehensive knowledge about enzymes, structure of substrates and interaction between them is required for successful

exploitation in biotechnological processes. Thus, intensive research is needed for studies like substrate specificity and specific activities, kinetic parameters, protein structure and stability.

3.1 Carbohydrate active enzymes

Many different enzymatic activities are involved in the depolymerization of plant biomass. Nature employs a large diversity of enzymes to act on the various substrates differing in monosaccharide composition, glycosidic linkages, side chain decorations [51]. The term Carbohydrate-Active enzymes, or CAZymes, refers to enzymes that act on different carbohydrate substrates, including their synthesis, breakdown and modifications as coined by the curators of the Carbohydrate-Active enZymes (CAZy) database (http://www.cazy.org/). CAZymes are classified in families based on amino acid sequence similarities, and these families largely reflect conserved three dimensional structural folds and stereochemical enzymatic mechanisms [52]. Up to date, the CAZy database includes more than 300 protein families to date divided into five enzyme classes and one non-catalytic class – carbohydrate-binding modules (CBMs). Table 3 presents the main CAZymes involved in cellulose, xyloglucan and starch degradation.

Table 3. Major Carbohydrate-Active enZyme (CAZy) classes involved in depolymerization of cellulose, xyloglucan and starch

SUBSTRATE	Enzyme class	CAZy families[a]	Function
Cellulose	endo-β-1,4-glucanase	GH5, 6, 7, 8, 9, 12, 44, 45, 48, 51, 74 124, 131	catalyze random cleavage of cellulose internal bonds at amorphous region
	cellobiohydrolase (CBH)	GH6, 7, 9, 48	exo-acting mode at the chain ends (CBHI at reducing end and CBHII at non-reducing ends)
	β-1,4-glucosidase	GH1, 3	hydrolyze cellooligosaccharides and cellobiose to glucose
	cellobiose/ cellooligosaccharide phosphorylase	GH94	phosphorolysis of β-glycosidic bonds in cellobiose and cellooligosaccharides
	lytic polysaccharide monooxygenase (LPMO)	AA9, AA10	cleavage of cellulose through an oxidative reaction mechanism
Xyloglucan	xyloglucan-specific endo-β-1,4-glucanase	GH5, 9, 12, 16, 44, 74	hydrolyze endo-β-1,4-glucosidic linkages in xyloglucan
	α-arabinofuranosidase	GH3, 43, 51, 54, 62	release of arabinose from the polymer
	α-xylosidase	GH31	removal of α-1,6-xylosyl residues from the non-reducing-end
	α-fucosidase	GH29, 95	cleavage of the fucosyle side chain
	α-galactosidase	GH27, 36	hydrolysis of terminal α-galactosyl moieties
	β-1,4-galactosidase	GH2, 35, 42	hydrolysis of β-galactosidases into monosaccharides
	β-1,4-glucosidase	GH1, 3	hydrolysis of terminal, non-reducing β-glucosyl residues with release of β-glucose
	lytic polysaccharide monooxygenase (LPMO)	AA9	cleavage of hemicellulose through an oxidative reaction mechanism
Starch	α-amylase	GH13, 57	acts on internal α-1,4-glycosidic bonds present in the amylose and amylopectin chains
	glucoamylase	GH15	breakdown of both α-1,4 and α-1,6 bonds from the non-reducing ends of amylase and amylopectin to release glucose
	α-1,4-glucosidase	GH31	release α-glucose from the non-reducing end of α-glucan oligomers or starch
	β-amylase	GH14	hydrolysis of α-1,4-glycosidic bonds from the non-reducing ends to remove maltose units
	lytic polysaccharide monooxygenase (LPMO)	AA13	cleavage of starch through an oxidative reaction mechanism

Based on CAZy database.

[a] GH – glycoside hydrolases, AA – auxiliary activities.

Most of the plant degrading enzymes belong to the biggest and best biochemically characterized glycoside hydrolase (GH) category of enzymes responsible for glycosidic bond cleavage [53]. Due to mechanistic and sequence similarities, this category also harbors some non-hydrolytic enzymes, like transglycosylases and phosphorylases. Approximately 40% of GH families contain plant cell wall deconstructing enzymes [54]. Many GHs have a modular organization with modules that retain their function and three dimensional fold independently. The most common of these are catalytic modules (CMs) and ancillary non-catalytic carbohydrate-binding modules (CBMs). While the CM function is to catalyze cleavage of substrate, CBMs typically bind to soluble and crystalline substrates and act like anchors that direct cognate enzymes to their target. CBMs enhance the catalytic activity through enhanced substrate binding or extension of substrate binding subsites if they are aligned with their respective CMs [55–57].

Other classes of enzyme activities are the glycosyltransferases (GTs), carbohydrate esterases (CEs) and polysaccharide lyases (PLs). The GTs are responsible for the biosynthesis of glycosidic bonds from phosphor-activated sugar donors [58]. The CEs assist GHs by removing esterified decorations, as several types of carbohydrates are substituted with acetyl moieties. The PLs represent a more uncommon activity that cleave e.g. uronic acid-containing polysaccharides via a β-elimination mechanism [59,60].

Due to the relatively new discovery of a novel oxidative mechanism for polysaccharide cleavage [61], the CAZy database was expanded with a new class of enzymes, auxiliary activities (AAs) [62]. Several AA families harbor catalytic activities that are involved in carbohydrate degradation through their ability to act in concert with GHs, PLs, and CEs enzymes to degrade carbohydrates [62]. This enzyme class also includes oxidative enzymes acting on mono- and oligosaccharides as well as lignin. The term lytic polysaccharide monooxygenase (LPMO) has been recently proposed for enzymes that catalyze oxidative cleavage of glycosidic linkages [63] and recently this class of enzymes has been classified into new CAZy families as AAs [62]. Since the main focus of this thesis is the discovery of new LPMOs, and as such, these enzymes are discussed in detail below.

3.2 Lytic polysaccharide monooxygenases

Recently, the conventional hydrolytic paradigm of plant biomass degradation has been challenged by several reports about the occurrence and functional importance of a novel class of recalcitrant polysaccharide degraders – lytic polysaccharide monooxygenases (LPMOs) [61,64,65]. Many new discoveries related to the structure, function, and diversity of these enzymes have been reported in the past few years.

3.2.1 Oxidative cleavage of glycosidic bonds: discovery of LPMOs

One of the first hints about oxidative enzymes influence on biomass degradation appeared already in 1974, where Eriksson et al. stated that not only hydrolytic but also oxidative enzymes are involved in cellulose degradation [66]. Almost for 40 years this notion has been largely ignored, until 2010 when proteins previously assigned into CBM33 and GH61 were reported to be oxidative enzymes and were reassigned as LPMOs [61,65,67].

In 1997, a report about the chitin-binding protein (CBP21) from the bacterium *Serratia marcescens* (formerly classified as CBM33 and now as AA10 enzyme) was published. This study demonstrated that CBP21 had binding affinity for chitin, but did not possess typical chitinase activity [68]. At the time, it was thought to be a non-catalytic protein, but remarkably it represented one of the main proteins in several secretome studies of *S. marcescens* grown on chitin [69]. Some years later, Vaaje-Kolstad et al. obtained the first crystal structure of CBP21 followed by studies that showed CBP21 being capable of enhancing the activity of chitinases [70–72]. These data indicated that CBP21 was more than a protein just capable of binding chitin. Indeed, in 2010, Vaaje-Kolstad et al. reported that CBP21 actually was an oxidative enzyme capable of cleaving the glycosidic bonds of the chitin chains [61]. A new enzyme class was identified, the LPMOs, later being classified in the auxiliary activity families of CAZy.

AA9 family enzymes were originally assigned in GH61 based mainly on a report showing weak endoglucanase activity of Cel61A from the fungus *Trichoderma reesei* [73]. The first indication that GH61 were not hydrolases came only in 2008 from the first crystal structure of Cel61B from *T. reesei*, which suggested a functional link between GH61 and CBM33 families [74] and no structural features typical of a GH. Transcriptomics and secretomics data demonstrated the widespread occurrence of GH61 proteins [75–79]. In 2010, Harris et al. reported the biochemical characterization of a GH61 enzyme from the fungus *Thielavia terrestris* and showed that it could enhance cellulase activity [80]. The AA9 boosting effect of cellulase activity was observed in several

reports and these enzymes are currently used as components in commercial enzyme preparations, *e.g*. Cellic CTec3 produced by Novozymes A/S [81–86].

Since 2012, previously classified proteins as CBM33 and GH61 are referred to as lytic polysaccharide monooxygenases (LPMOs) [63]. Moreover, in 2013 LPMOs were defined in CAZy database as auxiliary activity (AA) families [62].

3.2.2 Diversity and biological role of LPMOs

Lytic polysaccharide monooxygenases are currently found in auxiliary activity (AA) families 9, 10, 11, and 13 in the CAZy database [62]. At the moment, AA9 family consists only of fungal proteins and more than 300 putative AA9 sequences are reported in the CAZy database (Figure 6). AA10 family proteins are found in several domains in life, but more than 90% out of 1753 reported so far are putative bacterial sequences (Figure 6). The AA11 and AA13 families were discovered very recently and currently harbor only a small number of proteins and very few of them are functionally characterized [87,88]. There is a limited number of characterized enzymes in all LPMO AA families compared to the number of their putative sequences in CAZy database.

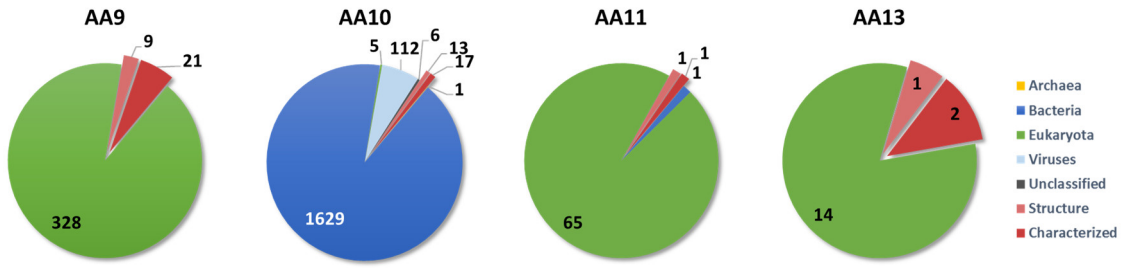


Figure 6. Classification and categories of LPMOs protein sequences in CAZy database.

LPMOs are broadly spread in both the bacterial and fungal domains, with the exception of AA10 which contains a significant number of viral sequences (Figure 6). It is unclear whether bacterial and fungal LPMOs evolved from a common ancestor, or whether their occurrence in such diverse species is a result of convergent parallel evolution. The genomes of plant biomass degrading fungi show a high number of genes encoding LPMOs and secretome analysis confirms that a subset of LPMOs are highly expressed when the fungi are grown on substrates containing recalcitrant carbohydrates [75,76,89–92]. Interestingly, some fungi possess more than 30 different AA9 encoding genes in their genomes [80], indicating a potentially important role(s) of the LPMOs in the lifestyle of the organism. The presence of genes encoding LPMOs in these fungi is correlated with

counterparts encoding hydrolytic enzymes. For example, in comparative analysis of 31 fungal genomes, more genes encoding LPMOs were identified than endo-β-1,4-glucanases from glycoside hydrolase families 5, 6 and 7 in total [93]. Representatives of AA9s are mainly found in filamentous Ascomycetes and Basidiomycetes [93,94].

The genomes of some bacteria typically encode one or two LPMOs, and in some cases even seven AA10 genes were detected [95]. Putative proteins from AA10 are mostly encountered in Proteobacteria, Firmicutes and Actinobacteria. There is an emerging literature implicating LPMOs with pathogenicity and virulence [96,97].

The production of LPMOs with a high degree of conservation across taxa suggests that different specificities of these enzymes are necessary for efficient plant biomass depolymerization. Families AA9 and AA13 appear to be highly related to plant biomass degradation and are the main focus of this PhD project.

3.2.3 Specificities and substrates for LPMOs

The first LPMO activity was demonstrated for the AA10 enzyme *Sm*LPMO10A (CBP21) from *S. marcescens* that showed oxidative cleavage of chitin, yielding chito-oligosaccharides with an aldonic acid reducing end moiety [61]. Then followed discoveries on cellulose degradation by both AA9 from *Thermoascus aurantiacus* (*Ta*LPMO9A, TaGH61) and AA10 from *Streptomyces coelicolor* A3(2) (*Sc*LPMO10C, CelS2) [67,98]. Subsequent work showed that LPMOs from both these families catalyze oxidative cleavage of chitin and cellulose [63,81,99,100]. It was shown that a pair of AA10s from *S. coelicolor* and *Thermobifida fusca*, that target different regions of the cellulose crystal, act in synergy, suggesting that organisms might need an assembly of LPMOs to degrade biomass [101]. Degradation of plant biomass by cellulases combined with AAs has been studied to only a limited extent. It has been shown that AA9 together with cellulases have a boosting effect on biomass degradation [80], and similar results have been shown for AA10 LPMOs in a study by Forsberg et al. (Forsberg et al 2011). The study of AA10 activity on plant biomass was conducted only in 2015, showing that the *Hc*AA10-2 produced by the marine bacterium *Hahella chejuensis* added to cellulases increased the release of reducing sugars from pretreated wheat straw [102]. Initially, it was assumed that LPMOs were only active on crystalline substrates. However, recent results have demonstrated that LPMO substrates also include soluble and semi-crystalline polysaccharides. The first example of an LPMO active on soluble substrates was *Nc*LPMO9C, an AA9 LPMO from *Neurospora crassa*. This enzyme was shown to be active on both crystalline cellulose and soluble cello-oligosaccharides hence expanding the known LPMO substrates to soluble polymers [103].

*Nc*LPMO9C has also been shown to be active on hemicelluloses that contain β-1,4-linked glucose units, such as xyloglucan, glucomannans, and β-glucan [64]. A similar activity profile has also been shown for AA9 LPMOs from *Podospora anserina* [104,105]. *Mt*LPMO9A, an AA9 enzyme from *Myceliophthora thermophila*, is the only LPMO so far demonstrated to have activity towards a hemicellulose with a non-glucose backbone as thus enzyme is able to cleave xylan chains [106]. It should be noted that this activity only was observed when xylan was mixed with cellulose (i.e. as a xylan-cellulose composite).

The substrate for the only one functionally characterized LPMO from AA11 family (*Ao*LPMO11 from *Aspergillus oryzae*) is chitin [87]. The newest family AA13 has only two characterized enzymes (*An*AA13 from *Aspergillus nidulans* and *Nc*AA13 from *N. crassa*) that showed lytic oxidation of starch [88,107].

LPMO substrates appear far more diverse than first anticipated and most likely more substrates for LPMOs will be revealed in near future.

3.2.4 Structure and mechanisms of LPMOs

The first LPMO structure was reported even before these enzymes were known as oxidative enzymes [70]. The structure of CBP21 (CBM33 at that time, now AA10) revealed that the majority of the conserved aromatic amino acids are located internally, opposite to what is typical for CBMs [57]. Determination of the first GH61 structure revealed similarities to CBP21 [74]. As previously mentioned, this gave the initial indication that GH61 were not glycoside hydrolases as it did not have active site groove or tunnel with acid/base catalytic residues found in GHs. Instead, these enzymes had a flat surface with a metal-binding site. The following structural studies of LPMOs brought more insights into LPMO function, explicitly with respect to the identification of active site residues, substrate binding interactions, internal electron transfer, and regioselectivity [67,80,99,108–111]. More details of the LPMO biochemical properties are presented below.

3.2.4.1 Tertiary structure and the binding site

There are 24 LPMO tertiary structures determined to date, nine representing AA9 enzymes and thirteen representing AA10 enzymes. In 2014, a structure of family AA11 enzyme was reported [87] and in 2015 the first structure from the newest family AA13 was published [107]. LPMOs share a high degree of structural similarity (Figure 7), despite lower than 10% sequence identity [107,112]. The most conserved residues are at or near the active-site surface. The overall fold of LPMOs has

been described as a small immunoglobulin-like structure that consists of a β-sandwich consisting of two β-sheets of typically 4–5 β-strands each [112,113]. The strands are linked by loops with a variable number of α-helix insertions [67,74,80,87,99,108,109]. The longer loops vary in size and conformation, particularly in segments that form the substrate-binding site, suggesting they may be important for substrate recognition [101,108,109,114].

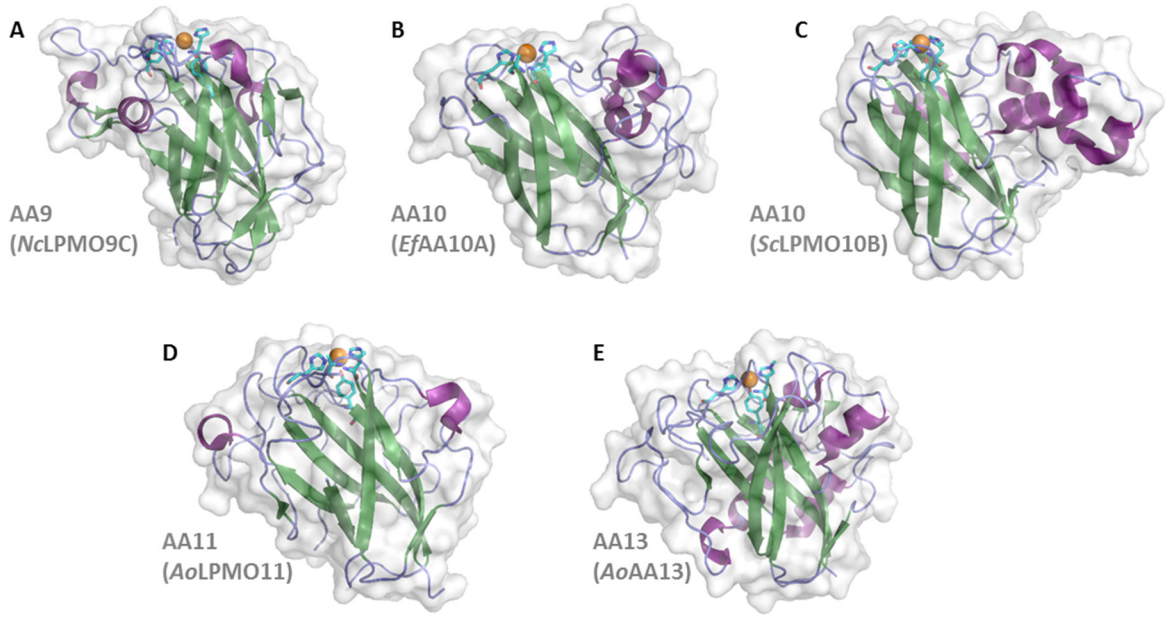


Figure 7. Representation of three dimensional structures from each LPMO CAZy family. (A) *Neurospora crassa* AA9 active both on cellulose and on hemicelluloses (PDB: 2D7U) [115]; (B) *Enterococcus faecalis* AA10 active on chitin (PDB: 4ALC) [116]; (C) *Streptomyces coelicolor* AA10 active on cellulose (PDB: 4OY6) [101]; (D) *Aspergillus oryzae* AA11 active on chitin (PDB: 4MAI) [87]; (E) *A. oryzae* AA13 active on starch (PDB: 4OPB) [107]. Overall structures are colored according to secondary structure and a semi-transparent surface is shown to depict topological details. Histidine and aromatic residues at the active site are shown as cyan sticks. The copper ions refined at the catalytic center are shown as orange spheres. The structures were rendered using PyMOL (The PyMOL Molecular Graphics System, version 1.3 Schrödinger, LLC).

The LPMO active site is positioned at the center or side of a flat surface area, which represents the substrate binding surface (Figure 7A–D). In contrast to the flatter binding surfaces seen in AA9, AA10 and AA11 enzymes, AA13 enzyme displays a shallow groove along the protein surface with the copper active site at the bottom of this cleft (Figure 7E). These differences in active site architecture provide clues on molecular recognition and interactions with different substrates. For example, the AA13 surface most likely accommodates better binding of helical structures that starch substrates form in amylose and amylopectin, as compared to the interaction between AA9–

AA11 LPMOs with the flat surfaces of major β-linked polysaccharides like cellulose and chitin. Unlike AA9s, the surface of AA11 and AA10 enzymes has few aromatic residues (usually only one or two), with the active site of the latter family being flanked with residues capable of forming hydrogen bonds to a potential polysaccharide substrates [87]. Aromatic residues on the AA9 surface surrounding the active site are commonly involved in enzyme-substrate interactions and the binding surface of LPMOs possesses two or three conserved tyrosine residues with rings parallel to the binding surface that suggest a role in substrate binding analogously to carbohydrate-binding modules [57,108]. The architecture of the LPMO catalytic domains fits very well with its activity: the core that possibly guides electrons to the copper center, with its conserved residues for electron transfer, and the loops that guide substrate to the copper center, with their substrate recognition motifs.

In some LPMOs, the catalytic domain is attached via a flexible linker to one or more carbohydrate-binding modules (CBM), suggesting that LPMOs target specific substrates via their appended CBMs. In contrast to the other AA families, family AA10 LPMOs show a large diversity of appended CBMs [63]. Some of the most common CBMs attached to family 10 LPMOs are CBM5, -12 and -14 that have putative chitin binding function and CBM2 that has putative cellulose binding function. As noted, family AA9 LPMOs show less diverse modular topology, with almost 75% of the enzymes registered in the CAZy database occurring as single domain proteins and approximately 20% having a CBM1 (cellulose binding module) at their C-terminus. The remaining AA9s have additional modules of unknown functions. Bennati-Granier et al. showed in their study that AA9 with the CBM1 released more oxidized products from substrate than without CBM1 [105]. The AA11 family has LPMO modules coupled to X278 domains of unknown function, but that are likely to be chitin binding given the activity so far shown for this family. The AA13 family shows LPMO modules attached to CBM20s, which are known to bind starchy substrates. For both the AA11 and AA13 families, the carbohydrate binding module was useful in the discovery of new activities of LPMO using “module-walking” approach, assuming a substrate binding module coupled to a domain of unknown function might lead the researcher to a novel enzyme family [107].

CBMs target their appended catalytic domain to the insoluble substrate, increase the enzyme local concentration and enhance the enzyme activity against insoluble substrates [57,117,118]. Thus, CBMs appended to LPMOs are an essential feature that needs to be taken into account when considering new enzyme cocktails for biomass degradation. Paper II gives insight into AA13 family LPMOs binding to substrates.

3.2.4.2 The copper active site

At the molecular level, LPMOs employ complex metal–oxygen chemistry [119]. Quinlan et al. in their X-ray crystallography studies were first to establish the copper identity of the metal center at the active site of LPMOs, as before it was assumed that variable metal ions could be accommodated to employ various metal ions [67]. The active site contains a mononuclear copper ion coordinated by the "histidine brace", arrangement that uses three nitrogen ligands contributed by the α-amino group and the sidechain of the N-terminal histidine together with a distal histidine sidechain to coordinate the copper in a T-shaped coordination plane (Figure 8) [67,119,120]. Enzymes use this single copper center to activate oxygen to generate reactive intermediates [121]. The histidine brace, a conserved feature in all characterized LPMOs, has come to be a defining characteristic of LPMOs (Figure 8–9) [110,122,123].

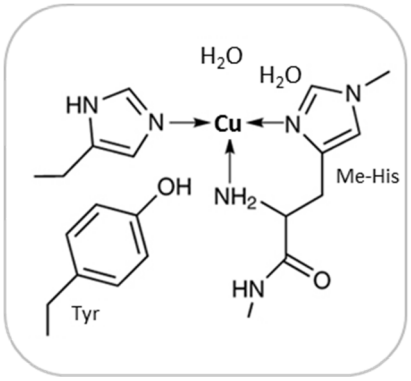


Figure 8. Schematic representation of active site with the histidine brace in AA9 LPMO. N-terminal histidine side chain is *N*-methylated. Adapted from [67,87].

In some structures there is still some uncertainty about the copper oxidation state and coordinating ligands, except for the three nitrogen ligands near the copper ion. Using electron paramagnetic resonance (EPR) spectroscopy, Quinlan et al. showed that AA9s contain a type II copper center [67]. Subsequent analyses of the active site demonstrated that there are differences between different LPMO families (Figure 9). Chitin-active bacterial AA10 enzymes (Figure 9B) differ from all other LPMOs as they display an EPR spectra that appears to be a mixture of a type II and type I copper signal [111,124]. However, there seems to be no consensus in the substrate specificities and the EPR signature, as a recent paper by Forsberg et al. shows a chitin-active family AA10 LPMO that displays an EPR signal more reminiscent of what has typically been observed for the cellulose active LPMOs [125]. For some LPMOs, the axial position of the copper site is partially

occupied by highly conserved alanine or phenylalanine residue (Figure 9). The significance of this amino acid is unclear, but has been proposed to steer the direction of dioxygen to the copper in the first stage of catalysis [101,111]. In addition, LPMOs have a conserved Glu/Gln residue that most likely coordinates a water molecule in the copper-coordination sphere (Figure 9). AA11 and AA13 family enzymes display type II copper EPR spectra as AA9s and cellulose-active AA10s [87,88,107]. AA11 LPMOs (Figure 9D) share features of both AA9 and AA10, having an axial tyrosine (as in AA9, Figure 9A) and the alanine near the other axial position of the copper site (as in AA10, Figure 9B–C). AA13 enzyme also slightly differs, having a tyrosine in one axial position (Figure 9E) and a loop in place of the alanine near the other axial coordination site [107]. At the moment, it is not clear how the mechanism of LPMO reaction is affected by these architectural differences in the active site, although this is obviously a highly interesting facet of LPMO research.

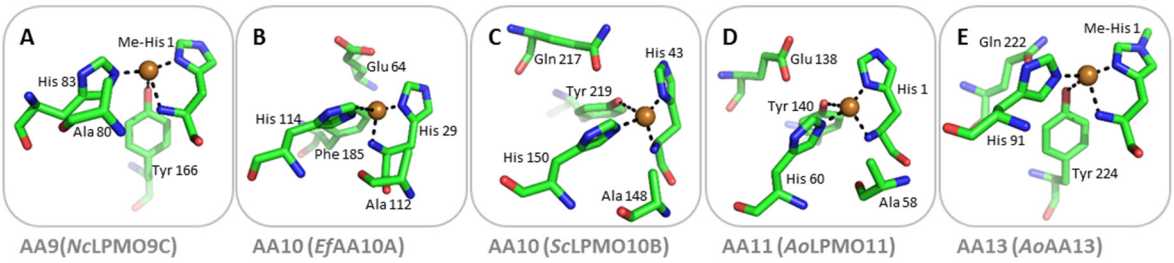


Figure 9. LPMO active sites. Active sites from each LPMO CAZy family (for more details see Figure 7) are shown in stick representation. Copper (golden sphere) coordinating groups are indicated by black dash lines. The conserved Glu/Gln residues, which are contributed from different regions of the enzymes around the copper center are also shown, as they are suggested to be important for defining the coordination geometry of the metals. The structures were rendered using PyMOL.

The N-terminal histidine side chain in the fungal LPMOs is post-translationally *N*-methylated when the enzymes are produced in their native organisms (i.e. filamentous fungi; [67,107,108]). Fungal LPMOs that are expressed in hosts that do not perform the methylation (like *Pichia pastoris* and *Escherichia coli*) retain activity [81,109,126], and therefore the role of this modification in LPMO activity remains unclear. It is not fully established whether bacterial LPMOs may have their *N*-terminal histidine post translationally methylated, but analysis of CBP21 produced by its native host, *S. marcescens*, was by mass spectroscopy shown not to be methylated [124].

3.2.4.3 Regioselectivity of oxidation

It is believed that oxidation by LPMOs leads to chain cleavage and introduction of new chain-ends, i.e. attachment points for GHs, in regions of the substrate otherwise inaccessible to these hydrolytic enzymes. This way the LPMOs act synergistically with GHs and boosts their activity [63,100]. The oxidation event can occur at either the C1 or the C4 carbon of the glycosidic bond, giving rise to a regioselectivity that varies in the LPMO enzyme family (Figure 10).

Studies have showed that the majority of fungal and bacterial LPMOs can be divided into three main groups: type 1, which oxidizes the C1 carbon of the glycosidic bond, type 2, which oxidizes the C4 carbon of the glycosidic bond, and type 3, which can oxidize either C1 or C4 [101,127]. Oxidation of the C1 carbon generates an aldonic acid at the downstream end of the polysaccharide chain, whereas oxidation of the C4 carbon yields a ketoaldose at the upstream end of the substrate (Figure 10) [120,127]. It should be noted that the aldonic acid resulting from C1 oxidation is in a pH dependent equilibrium with the corresponding δ-1,5-lactone (Figure 10) [61]. On the other hand, C4 oxidation product, the 4-ketoaldose, is in an equilibrium with its corresponding gemdiol (Figure 10). LPMOs are also capable of releasing native oligosaccharides. This can occur when oxidation occurs close to a chain end, i.e. close to the reducing end (downstream end) for C1 oxidizing enzymes and near the non-reducing end (upstream end) for the C4 oxidizing enzymes. Enzymes that show mixed activity (see below) can cause release of native oligosaccharides from intra chain cleavage events (i.e. when C1 oxidation occurs on the upstream end and C4 oxidation occurs close to downstream [101]).

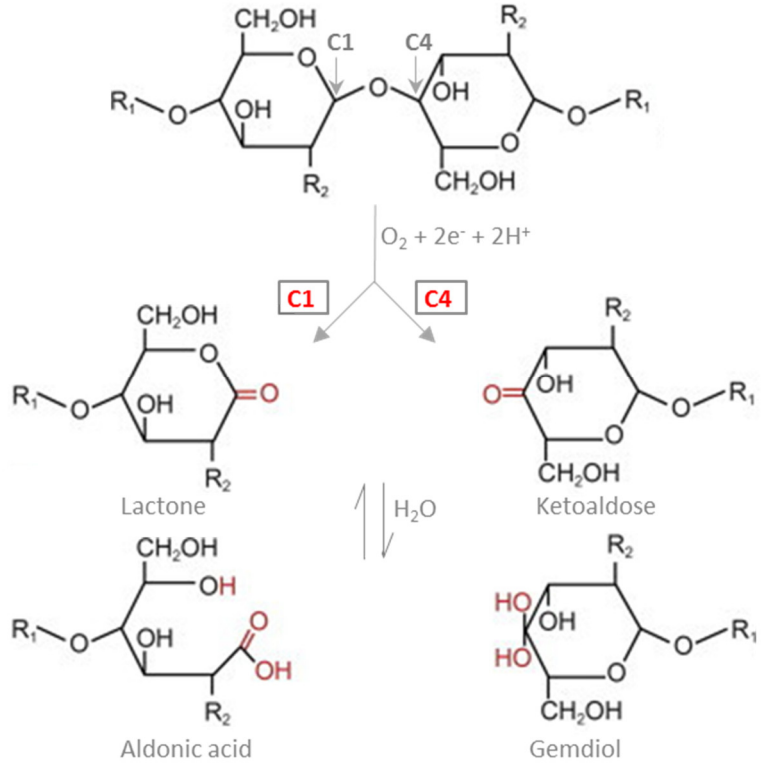


Figure 10. The LPMO reaction scheme. LPMOs catalyze oxidative cleavage at the C1 or C4 carbon within a polysaccharide chain leading to chain cleavage. Oxidation at C1 results in the formation of a lactone that spontaneously is hydrolyzed to aldonic acid. C4 oxidation leads to the formation of a ketoaldose and that is in equilibrium with its geminal diol form. R_1 represents polymer chain and R_2 represents -OH or -NHAc group in cellulose or chitin, respectively. Adapted from [97].

The first LPMO that showed oxidation of both C1 and C4 carbons was *Ta*LPMO9A [67]. Quinlan et al. also observed gemdiol products oxidized at C6 by *Ta*LPMO9A but C6 oxidation is a non-productive reaction pathway that does not lead to cleavage of the glycosidic bond in chitin or cellulose [65,67,127]. AA9s that selectively oxidize at C1, C4, or both C1 and C4 are currently described. Until recently, AA10s were only known to oxidize at C1 [61,98,124], but in 2014 Forsberg et al. identified AA10s that can oxidize cellulosic substrates at both C1 and C4, i.e. having mixed activity [101]. In LPMO structures, differences in the accessibility of the solvent-facing axial position associate with the regioselectivity [101,115]. AA9s that only oxidize C1 have a tyrosine residue that restricts access to this position and renders dioxygen activation more likely in the equatorial position, whereas this solvent-facing axial position is unrestricted in C4 oxidizing AA9s. AA9s that can oxidize at both C1 and C4 are in an intermediate situation, and the tyrosine is substituted with a proline. In C1 oxidizing AA10s, the same function is achieved by a conserved alanine (Figure 9B–C) [101,113]. The newest families AA11 and AA13 enzymes are currently known to oxidize only at

C1, but presumably new members with different regioselectivities may emerge [87,88,107]. Understanding the chemical basis for differences in regioselectivity and the synergy between LPMOs with different selectivities may help in developing new enzyme cocktails for biomass conversion. Indeed, synergy between a strict C1 oxidizing LPMO and an LPMO with mixed C1/C4 activity has indeed already been reported [101].

3.2.4.4 Reaction mechanism

Since the discovery of LPMOs in 2010, it was known that these enzymes utilize molecular oxygen to cleave glycosidic bonds [61]. Experiments with isotope-labeling ($H_2{}^{18}O$ and ${}^{18}O_2$) proved that LPMOs incorporate one oxygen atom from molecular oxygen in their reaction products, which confirms the need of O_2 in catalysis. In addition to their copper cofactor and O_2, a redox-active cofactor with metal reducing capacity is required for LPMO activity [61]. The reaction mechanism itself is largely unknown, but some studies have been conducted to shed light on it, although conclusions so far not have been coherent. Kim et al. suggested that the mechanism involves the formation of a copper-oxyl radical that abstracts a hydrogen atom from one of the carbons, which results in substrate hydroxylation via an oxygen-rebound mechanism, followed by elimination to cleave the glycosidic bond and incorporation of a single oxygen atom from molecular oxygen into the products [119]. Kjaergaard et al., on the other hand, performed spectroscopic experiments, leading them to suggest a mechanism where superoxide is the reactive oxygen species involved in the initial hydrogen abstraction event [122]. Thus, there is a need for more research to clarify the various steps in the reaction mechanism.

As previously mentioned, the LPMO needs to obtain an electron from an external source in order to work. Several small reducing compounds, such as ascorbic acid, L-cysteine, reduced glutathione, have been used *in vitro* [81,98,107]. In addition, it has been shown that the electrons for fungal LPMOs can be supplied by cellobiose dehydrogenase (CDH), a multi-domain enzyme only secreted by some fungi [65,84]. CDHs catalyze the oxidation of cellobiose, longer cellodextrins or other oligosaccharides to corresponding lactones that are spontaneously converted to the aldonic acids. The oxidation event leads to the reduction of FAD to $FADH_2$ in the cytochrome domain of CDH that subsequently transfers electrons to a heme domain appended to the cytochrome domain. The heme domain acts as a mediator and transfers electrons to the active site of LPMOs, which thereby can activate oxygen [65,84,128,129]. Several studies have shown that AA9s and CDHs often are co-expressed [89,90,93], indicating that the CDH and the LPMOs possibly work together. Also, CDH gene knockouts have been shown to result in loss of cellulolytic activity compared with wild-

type in *N. crassa* [65]. However, CDH is not produced by all fungi that encode LPMOs and to date no bacterial CDHs have been identified, leading to the assumption that other (enzymatic or chemical) natural electron sources exist. Indeed, a study on the cellulolytic machinery of *Cellvibrio japonicus* showed that an enzyme with a redox function is important for the growth of the bacterium on cellulose [130]. This enzyme may represent the bacterial analog of CDH. Microorganisms may also utilize other approaches for providing electron donors. Other potential electron sources can be found in the plant cell wall, as lignin, possible hemicelluloses, or gallic acid [80,100,131,132]. However, very little is known about the source of reducing cofactors within different type of plant biomass degrading organisms and the influence of these sources on LPMO efficiency *in vivo*.

4 Fungal degradation of plant biomass

In nature, filamentous fungi and bacteria are the major degraders of plant biomass, and this process has high significance in the global carbon cycle [133–135]. The enzymatic degradation of complex plant polysaccharides is achieved through the evolution of secreted enzyme systems consisting of multiple carbohydrate-active enzymes, typically acting together as a cocktail with complementary, synergistic activities and modes of action [27,136]. A more profound understanding of microbial degradation of plant biomass and the cost-effective production of enzymes that deconstruct the complex and heterogeneous chemical linkages, that are present in this type of substrates, could help to improve the feasibility of biotechnological processes that are envisioned to contribute to a modern bioeconomy. Using microorganisms as biotechnological sources for industrially relevant enzymes has stimulated interest in the exploration of extracellular enzymatic activities in several microorganisms.

Filamentous fungi are well known for their ability to secrete impressive levels of polysaccharide degrading enzymes due to their metabolic reliance on this type of complex substrates in the natural ecological niches [134]. The deconstruction of macromolecular or supramolecular substrates to simple monomers or oligomers allows the internalization of these simpler components by specialized transporters in contrast to larger substrates that cannot be transported into the cells [137]. There is likely a large unexploited potential in the rich diversity of enzymes produced by fungi, which offers a formidable toolkit for the design of tailor-made efficient enzyme cocktails to target different substrates. Recently it has been shown that there are far better fungi for saccharification of plant biomass than a model fungus *Trichoderma reesei* that has been used in industry for a long time. For example, Navaro et al. in their study introduced a *Laetisaria arvalis* strain that has 7.5-fold higher cellulolytic activity than a *T. reesei* industrial strain [138]. Interestingly, more than half of the current commercial enzymes are of fungal origin and mainly from the *Aspergillus* genus (Figure 11). Fungal enzymes are frequently used in various industrial applications like food, animal feed, textiles, detergents, biofuels and biochemicals [139,140]. The demand for new industrial enzymes is on a continuous rise driven by a growing need for sustainable solutions.

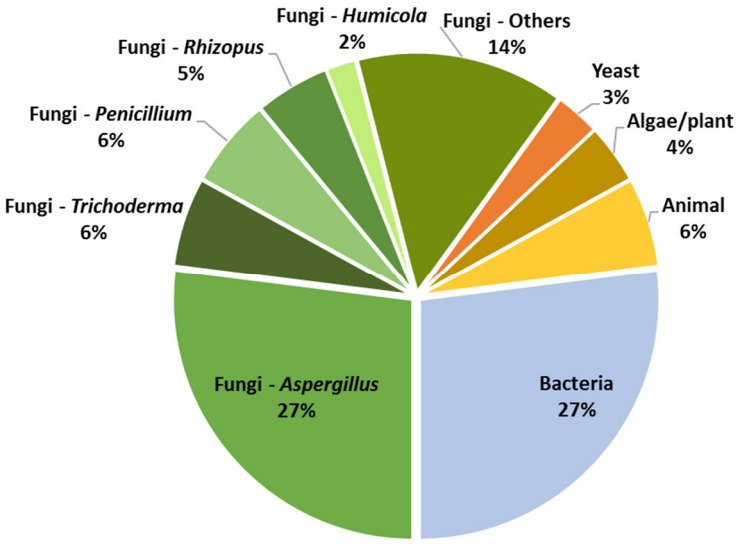


Figure 11. Distribution of industrial enzymes by origin of donor organism based on data from AMFEP (2009) with specification of the top five fungal organisms [139].

So far over 250 fungal genomes have been sequenced [141]. Moreover, the rate of fungal genome sequencing has been increasing substantially in recent years; as an example is a massive "1000 Fungal Genomes" project [142]. These fungal genomes are revealing a large number of putative carbohydrate active enzymes. *In silico* annotations provide reasonable level of information, but experimental analyses remain necessary to identify and characterize novel proteins and for better knowledge on the enzyme batteries secreted in response to distinct glycan structures. To design efficient enzyme cocktails and to achieve maximal synergy, the doses of enzymes as well as their optimal relative ratios need to be considered.

In this thesis, the composition of enzymes produced for starch degradation was studied by employing secretome analysis of *Aspergillus nidulans* together with amylolytic biochemical assays (Paper I), whereas the binding properties of the starch active LPMO to starch and the starch mimic β-cyclodextrin were studied in Paper II. The action of a novel LPMO from the plant pathogenic *Fusarium graminearum* on cellulose and hemicellulose was also demonstrated (Paper III). Therefore, a survey of proteomic studies and these two filamentous fungi, *A. nidulans* and *F. graminearum*, will be described in more detail in this section.

4.1 Secretomics: a powerful tool in analysis and discovery of novel CAZymes

The availability of sequenced genomes together with the progress in resolution of mass spectrometry (MS) analysis led to a burst numbers of fungal proteome studies. Proteomics has grown to a powerful tool, as it generates valuable insight on the actual production and secretion of proteins rather than the potential to do that, which is accessed through genomic analyses. Proteomics can be used to prospect new organisms and novel enzymes, and to reveal strategies used by different fungi for plant biomass degradation. This technique provides information at the protein level, and allows the determination of the composition of proteins produced by these fungi under specific growth conditions [143]. The integration of such data together with biochemical and structural data is likely to be an important component in the design of enzyme mixtures to optimize the breakdown of recalcitrant substrates with specific biomass composition.

Enzymes produced by filamentous fungi for polysaccharide degradation are mainly released to the extracellular medium. The type of proteomics used to study all the secreted proteins is called secretomics. Due to the development of fungal genome projects, theoretical secretomes can be predicted *in silico*, based on the identification of signal peptide sequences in putative proteins [144]. However, these prediction methods have a variety of issues as there is a high likelihood for false positives and false negatives [145,146]. Moreover, the actual secretome is highly dependent gene regulatory networks that depend on environmental conditions, including the substrate, temperature, pH and growth phase. Therefore, proteomic approaches are necessary to reveal these various aspects, which are impossible to predict from the *in silico* analysis.

Secretomics studies of filamentous fungi frequently identify a variety of fungal carbohydrate-active enzymes (see section 3.1) involved in plant biomass degradation [89,138,147–149]. Several fungi secrete not only hydrolytic enzymes but also an assembly of oxidative enzymes, like cellobiose dehydrogenase (CDH that is classified in AA3 and AA8) and LPMOs [138,147,150], that also have been associated with polysaccharide degradation. For instance, the expression of 10 out of 14 predicted LPMOs was induced during the ascomycete *N. crassa* growth on cellulose [75], and 7 out of 16 in the basidiomycete *Laetisaria arvalis* [138]. Interestingly, the occurrence of genes encoding LPMOs in ascomycetes and basidiomycetes is strongly associated with the presence of genes encoding hydrolytic enzymes, and the lack of LPMOs translates to the absence of CDH [112].

While considerable focus has been imparted on lignocellulose matrices [12], much less is known about the enzyme cocktails deployed by fungi for the degradation of starch, both with respect to their class, composition and distribution. Moreover, the abundance and impact of LPMOs on starch

degradation by fungi are currently unclear. The study in Paper I provides an overview into the significant capabilities of the ascomycete *A. nidulans* to degrade starch.

4.2 *Aspergillus nidulans*

The genus *Aspergillus* includes a diverse group of filamentous fungi that consists of some pathogenic species (e.g. *A. fumigatus*, a serious human pathogen [151]) and some that are very beneficial (e.g. *A. niger*, an industrial citric acid producer [152]). *A. nidulans* has been used as a model organism for more than 50 years and is one of the first organisms to have its genome sequenced at the Broad Institute [153]. It is a close homologue to the industrially important *A. niger* and *A. oryzae* species (Galagan et al., 2005; Wortman et al., 2009). However, the number of proteomic studies on *Aspergilli* is relatively low and so far only 11% of the open reading frames (ORFs) have been verified as proteins in *A. nidulans* (Figure 12) [156]. Genome analysis provides little mechanistic insight into the extraordinary capabilities of this fungus to secrete proteins in response to growth on various complex plant derived carbon sources.

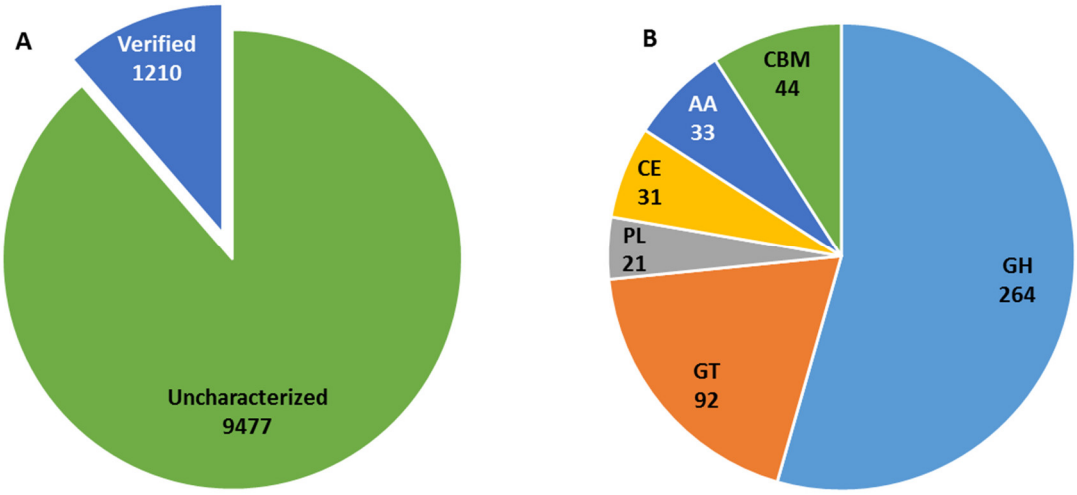


Figure 12. (A) Graphical view of protein coding genes in *Aspergillus nidulans* (modified from the *Aspergillus* Genome Database (AspGD; http://www.aspgd.org) as of Jan 31, 2016). (B) Distribution of CAZyme modules or domains in *Aspergillus nidulans*; AA, auxiliary activity; CBM, carbohydrate binding module; CE, carbohydrate esterase; GH, glycoside hydrolase; GT, glycosyltransferase; PL, polysaccharide lyase.

Proteomic analysis has been applied in some *Aspergillus* research areas. In a study assessing the contribution of *A. nidulans* for leaf litter degradation, secretome analysis revealed a total of 90 proteins, where 21 of them were related to cellulose degradation: four exoglucanases

(cellobiohydrolases), ten endoglucanases (in fact, five of them are now stated as LPMOs), and seven β-glucosidases [157]. In another case, the extracellular proteome of *A. nidulans* grown on sorghum stover was studied at different time points [149]. A total of 294 secreted proteins were identified including cellulases, hemicellulases, pectinases, chitinases and many others. This study shows that most of the enzymes were secreted already at day 1 of growth and changes in the relative abundance over time indicates that sets of secreted enzymes are "custom-made" depending on the substrate. Remarkably, CDH and family AA9 enzymes were found in high abundance. In another study, Couturier et al. identified a total of 66 proteins secreted by *A. nidulans* when grown on maize bran [158]. They identified seven hemicellulases (from GH10, GH11, GH39, GH43, GH62 and GH93 families) and five β-1,3-glucanases (from GH17, GH55 and GH81 families) but no β-1,4-endoglucanases or cellobiohydrolases. The difference in secreted enzyme cocktails of *A. nidulans* may be attributed to difference in composition of substrates and growth conditions. All these results confirm that *A. nidulans* is capable of utilizing a range of plant polysaccharide polymers by expressing a wide variety of enzymes which reflects the saprophytic lifestyle of this fungus.

There are no secretome studies to date about *A. nidulans* utilization of starch. Paper I displays results of secretome and biochemical assays of activities of enzymes expressed by this fungus on starches from different sources.

4.3 *Fusarium graminearum*

One of the areas of fungal research in which plant degradation has received attention is plant pathogenicity. Fungal plant pathogens can cause devastating crop losses. One of the examples is a well-known pathogenic ascomycete fungus *Fusarium graminearum* [159]. It causes substantial economic losses due to damage of high-yielding crops and their quality, and contaminating them with mycotoxins hazardous to humans and animals [160]. This fungus causes Fusarium head blight disease most commonly on wheat and barley, and Fusarium stalk and ear rot disease on maize [153,160]. *F. graminearum* even made it to the "Top 10 fungal pathogens" list [161]. Genomic research of *F. graminearum* has started in the late 90s but only in 2015 a fully completed genome sequence was released [162]. Remarkably, this pathogen has almost 600 predicted CAZymes (Figure 13A). Moreover, more than 10% of all CAZymes are AA family enzymes, where none of 18 putative LPMOs are yet characterized (Figure 13B). Analysis of genes coding CAZymes, including AA9s, showed that most of them are up-regulated during plant infection [163].

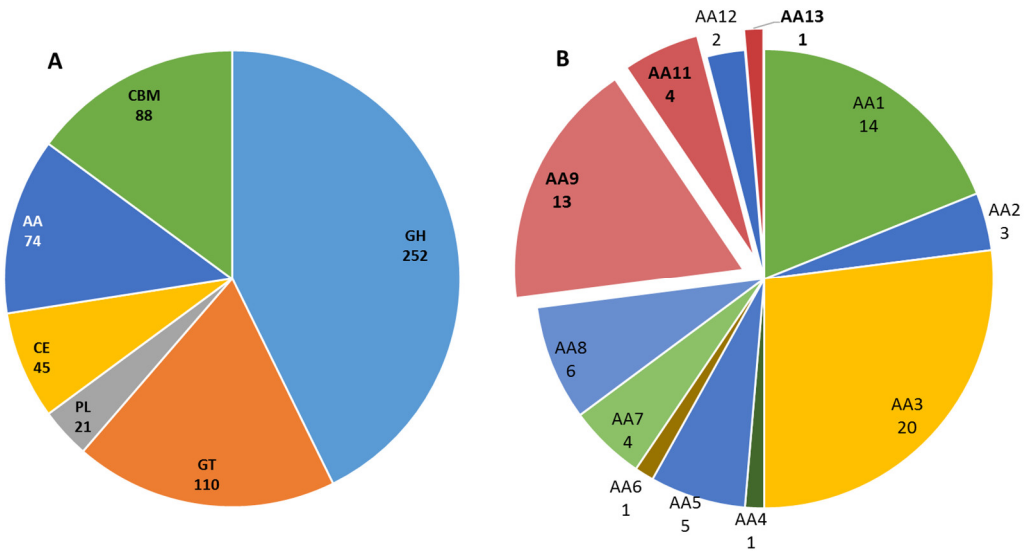


Figure 13. (A) Distribution of CAZyme modules or domains in *Fusarium graminearum*; AA, auxiliary activity; CBM, carbohydrate binding module; CE, carbohydrate esterase; GH, glycoside hydrolase; GT, glycosyltransferase; PL, polysaccharide lyase. (B) Distribution of auxiliary activity families in *Fusarium graminearum*. LPMOs are in red.

The proteomics approach can complement and extend gene expression studies on the fungus-host interaction. Paper et al. [164] in their proteomics study identified a total of 229 secreted proteins during growth on different carbon sources, including birchwood xylan, pectin, corn stover, maize bran and stover, dried distillers' grains and others. Intriguingly, only six proteins were commonly secreted in all media, whereas the most of extracellular proteins were formed as a response to the particular culture conditions. In another study, *F. graminearum* was grown on four lignocellulosic substrates (corn cobs, wheat bran, hop cell walls, and birchwood) [165]. In this case only one out of 57 cell wall-degrading enzymes (total of 180 unique proteins) was common to all cases. Only during growth on birchwood, two AA9s were secreted. Briefly, *F. graminearum* adapts itself to the biomass it has encountered by the secretion of a specific enzymatic cocktail.

Even though the genome of *F. graminearum* has been sequenced and the resulting assembly displays high quality, still a significant number of extracellular proteins are not functionally assigned. In Paper 3, the first LPMO from *F. graminearum* is characterized.

5 The aim of the present study

The work presented in this thesis has focused on fungal degradation of polysaccharides, with a special attention on lytic polysaccharide monooxygenases. The aim has been three-sided:

- Investigation of *Aspergillus nidulans* protein secretion when the fungus was grown on various starches.
- Analysis of functional properties of starch-specific LPMOs.
- Biochemical characterization of a novel LPMO from the fungus *Fusarium graminearum*.

References

1. Naik SN, Goud VV, Rout PK, Dalai AK. Production of first and second generation biofuels: A comprehensive review. *Renew Sustain Energy Rev*. 2010;14(2):578-597.
2. Scarlat N, Dallemand JF, Monforti-Ferrario F, Nita V. The role of biomass and bioenergy in a future bioeconomy: Policies and facts. *Environ Dev*. 2015;15:3-34.
3. Bentsen NS, Felby C. Biomass for energy in the European Union - a review of bioenergy resource assessments. *Biotechnol Biofuels*. 2012;5(1):25.
4. Smeets EMW, Faaij APC, Lewandowski IM, Turkenburg WC. A bottom-up assessment and review of global bio-energy potentials to 2050. *Prog Energy Combust Sci*. 2007;33(1):56-106.
5. Mohr A, Raman S. Lessons from first generation biofuels and implications for the sustainability appraisal of second generation biofuels. *Energy Policy*. 2013;63(100):114-122.
6. Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science*. 2007;315(5813):804-807.
7. Aro EM. From first generation biofuels to advanced solar biofuels. *Ambio*. 2016;45(1):24-31.
8. Balan V. Current challenges in commercially producing biofuels from lignocellulosic biomass. *ISRN Biotechnol*. 2014;2014:463074.
9. Escobar JC, Lora ES, Venturini OJ, Yáñez EE, Castillo EF, Almazan O. Biofuels: Environment, technology and food security. *Renew Sustain Energy Rev*. 2009;13(6-7):1275-1287.
10. Ho DP, Ngo HH, Guo W. A mini review on renewable sources for biofuel. *Bioresour Technol*. 2014;169:742-749.
11. Cannella D, Jørgensen H. Do new cellulolytic enzyme preparations affect the industrial strategies for high solids lignocellulosic ethanol production? *Biotechnol Bioeng*. 2014;111(1):59-68.
12. Bouws H, Wattenberg A, Zorn H. Fungal secretomes - Nature's toolbox for white biotechnology. *Appl Microbiol Biotechnol*. 2008;80(3):381-388.
13. Buckeridge MS, Pessoa dos Santos H, Tiné MAS. Mobilisation of storage cell wall polysaccharides in seeds. *Plant Physiol Biochem*. 2000;38(1-2):141-156.
14. Carpita NC, Gibeaut DM. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J*. 1993;3(1):1-30.
15. Cosgrove DJ. Growth of the plant cell wall. *Nat Rev Mol Cell Biol*. 2005;6(11):850-861.
16. Doblin MS, Pettolino F, Bacic A. Plant cell walls: the skeleton of the plant world. *Funct Plant*

Biol. 2010;37(5):357-381.

17. Doherty WOS, Mousavioun P, Fellows CM. Value-adding to cellulosic ethanol: Lignin polymers. *Ind Crops Prod*. 2011;33(2):259-276.
18. Schultink A, Liu L, Zhu L, Pauly M. Structural diversity and function of xyloglucan sidechain substituents. *Plants*. 2014;3:526-542.
19. Rose JKC, Bennett AB. Cooperative disassembly of the cellulose-xyloglucan network of plant cell walls: Parallels between cell expansion and fruit ripening. *Trends Plant Sci*. 1999;4(5):176-183.
20. Saha B. Lignocellulose biodegradation and applications in biotechnology. *ACS Symp Ser*. 2004;889:2-34.
21. Brown RM. Cellulose structure and biosynthesis: What is in store for the 21st century? *J Polym Sci Part a-Polymer Chem*. 2004;42(3):487-495.
22. Habibi Y, Lucia LA, Rojas OJ. Cellulose nanocrystals: Chemistry, self-assembly, and applications. *Chem Rev*. 2010;110(6):3479-3500.
23. Donaldson L. Cellulose microfibril aggregates and their size variation with cell wall type. *Wood Sci Technol*. 2007;41(5):443-460.
24. U.S. Department of Energy Office of Science. Genomics: GTL Roadmap: Systems biology for energy and environment. 2005:204.
25. Gibson LJ. The hierarchical structure and mechanics of plant materials. *J R Soc Interface*. 2012;9(76):2749-2766.
26. Fernandes AN, Thomas LH, Altaner CM, Callow P, Forsyth VT, Apperley DC, Kennedy CJ, Jarvis MC. Nanostructure of cellulose microfibrils in spruce wood. *Proc Natl Acad Sci U S A*. 2011;108(47):1195-1203.
27. Payne CM, Knott BC, Mayes HB, Hansson H, Himmel ME, Sandgren M, Ståhlberg J, Beckham GT. Fungal cellulases. *Chem Rev*. 2015;115(3):1308-1448.
28. Scheller HV, Ulvskov P. Hemicelluloses. *Annu Rev Plant Biol*. 2010;61(1):263-289.
29. Zabotina OA. Xyloglucan and its biosynthesis. *Front Plant Sci*. 2012;3:1-5.
30. Tuomivaara ST, Yaoi K, O'Neill MA, York WS. Generation and structural validation of a library of diverse xyloglucan-derived oligosaccharides, including an update on xyloglucan nomenclature. *Carbohydr Res*. 2015;402:56-66.
31. Hsieh YSY, Harris PJ. Xyloglucans of monocotyledons have diverse structures. *Mol Plant*. 2009;2(5):943-965.
32. Hayashi T, Kaida R. Functions of xyloglucan in plant cells. *Mol Plant*. 2011;4(1):17-24.

33. Vogel J. Unique aspects of the grass cell wall. *Curr Opin Plant Biol*. 2008;11(3):301-307.
34. Park YB, Cosgrove DJ. Xyloglucan and its interactions with other components of the growing cell wall. *Plant Cell Physiol*. 2015;56(2):180-194.
35. Ebringerova A, Hromadkova Z, Heinze T. Hemicellulose. *Adv Polym Sci*. 2005;186:1-67.
36. Burton RA, Fincher GB. Evolution and development of cell walls in cereal grains. *Front Plant Sci*. 2014;5:456.
37. Buckeridge MS. Seed cell wall storage polysaccharides: models to understand cell wall biosynthesis and degradation. *Plant Physiol*. 2010;154(3):1017-1023.
38. Buléon A, Colonna P, Planchot V, Ball S. Starch granules: structure and biosynthesis. *Int J Biol Macromol*. 1998;23(2):85-112.
39. Zeeman SC, Kossmann J, Smith AM. Starch: its metabolism, evolution, and biotechnological modification in plants. *Annu Rev Plant Biol*. 2010;61:209-234.
40. Pérez S, Bertoft E. The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review. *Starch/Staerke*. 2010;62(8):389-420.
41. Favaro L, Jooste T, Basaglia M, Rose SH, Saayman M, Görgens JF, Casella S, van Zyl WH. Designing industrial yeasts for the consolidated bioprocessing of starchy biomass to ethanol. *Bioengineered*. 2013;4(2):97-102.
42. Shibanuma K, Takeda Y, Hizukuri S, Shibata S. Molecular structures of some wheat starches. *Carbohydr Polym*. 1994;25(2):111-116.
43. Buléon A, Colonna P. Physicochemical behaviour of starch in food applications. In: *The Chemical Physics of Food*.; 2007:20-67.
44. Gallant DJ, Bouchet B, Baldwin PM. Microscopy of starch: evidence of a new level of granule organization. *Carbohydr Polym*. 1997;32(3-4):177-191.
45. Tang H, Mitsunaga T, Kawamura Y. Molecular arrangement in blocklets and starch granule architecture. *Carbohydr Polym*. 2006;63(4):555-560.
46. Angellier-Coussy H, Putaux JL, Molina-Boisseau S, Dufresne A, Bertoft E, Perez S. The molecular structure of waxy maize starch nanocrystals. *Carbohydr Res*. 2009;344(12):1558-1566.
47. Tester RF, Karkalas J, Qi X. Starch - composition, fine structure and architecture. *J Cereal Sci*. 2004;39(2):151-165.
48. Ragauskas AJ. The path forward for biofuels and biomaterials. *Science*. 2006;311(5760):484-489.

49. Cherubini F. The biorefinery concept: Using biomass instead of oil for producing energy and chemicals. *Energy Convers Manag*. 2010;51(7):1412-1421.
50. Jordan DB, Bowman MJ, Braker JD, Dien BS, Hector RE, Lee CC, Mertens JA, Wagschal K. Plant cell walls to ethanol. *Biochem J*. 2012;442:241-252.
51. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res*. 2009;37:D233-D238.
52. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2014;42(D1):490-495.
53. Henrissat B, Davies G. Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol*. 1997;7(5):637-644.
54. Gilbert HJ. The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiol*. 2010;153(2):444-455.
55. Hashimoto H. Recent structural studies of carbohydrate-binding modules. *Cell Mol Life Sci*. 2006;63(24):2954-2967.
56. Guillén D, Sánchez S, Rodríguez-Sanoja R. Carbohydrate-binding domains: multiplicity of biological roles. *Appl Microbiol Biotechnol*. 2010;85(5):1241-1249.
57. Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J*. 2004;382:769-781.
58. Henrissat B, Sulzenbacher G, Bourne Y. Glycosyltransferases, glycoside hydrolases: surprise, surprise! *Curr Opin Struct Biol*. 2008;18(5):527-533.
59. Garron ML, Cygler M. Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. *Glycobiology*. 2010;20(12):1547-1573.
60. Lombard V, Bernard T, Rancurel C, Brumer H, Coutinho PM, Henrissat B. A hierarchical classification of polysaccharide lyases for glycogenomics. *Biochem J*. 2010;432(3):437-444.
61. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørlie M, Eijsink VGH. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*. 2010;330(6001):219-222.
62. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels*. 2013;6(1):41.
63. Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VG. Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels*. 2012;5(1):45.

64. Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WGT, Ludwig R, Horn SJ, Eijsink VGH, Westereng B. Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci*. 2014;111(17):6287-6292.

65. Phillips CM, Beeson WT, Cate JH, Marletta MA. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem Biol*. 2011;6(12):1399-1406.

66. Eriksson KE, Pettersson B, Westermark U. Oxidation: An important enzyme reaction in fungal degradation of cellulose. *FEBS Lett*. 1974;49(2):282-285.

67. Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JCN, Johansen KS, Krogh KBRM, Jørgensen CI, Tovborg M, Anthonsen A, et al. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci*. 2011;108(37):15079-15084.

68. Watanabe T, Kimura K, Sumiya T, Nikaidou N, Suzuki K, Suzuki M, Taiyoji M, Ferrer S, Regue M. Genetic analysis of the chitinase system of *Serratia marcescens* 2170. *J Bacteriol*. 1997;179(22):7111-7117.

69. Suzuki K, Suzuki M, Taiyoji M, Nikaidou N, Watanabe T. Chitin binding protein (CBP21) in the culture supernatant of *Serratia marcescens* 2170. *Biosci Biotechnol Biochem*. 1998;62(1):128-135.

70. Vaaje-Kolstad G, Houston DR, Riemen AHK, Eijsink VGH, Van Aalten DMF. Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21. *J Biol Chem*. 2005;280(12):11313-11319.

71. Vaaje-Kolstad G, Horn SJ, Van Aalten DMF, Synstad B, Eijsink VGH. The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. *J Biol Chem*. 2005;280(31):28492-28497.

72. Eijsink VGH, Vaaje-Kolstad G, Vårum KM, Horn SJ. Towards new enzymes for biofuels: lessons from chitinase research. *Trends Biotechnol*. 2008;26(5):228-235.

73. Karlsson J, Saloheimo M, Siika-Aho M, Tenkanen M, Penttilä M, Tjerneld F. Homologous expression and characterization of Cel61A (EG IV) of *Trichoderma reesei*. *Fed Eur Biochem Soc J*. 2001;268(24):6498-6507.

74. Karkehabadi S, Hansson H, Kim S, Piens K, Mitchinson C, Sandgren M. The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 Å resolution. *J Mol Biol*. 2008;134(2):199-218.

75. Tian C, Beeson WT, Iavarone AT, Sun J, Marletta M a, Cate JHD, Glass NL. Systems analysis of plant cell wall degradation by the model filamentous fungus *Neurospora crassa*. *Proc Natl Acad Sci U S A*. 2009;106(52):22157-22162.

76. Eastwood DC, Floudas D, Binder M, Majcherczyk A, Schneider P, Aerts A, Asiegbu FO, Baker SE, Barry K, Bendiksby M, et al. The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. *Science*. 2011;333(6043):762-765.

77. Wymelenberg A Vanden, Gaskell J, Mozuch M, Sabat G, Ralph J, Skyba O, Mansfield SD, Blanchette RA, Martinez D, Grigoriev I, et al. Comparative transcriptome and secretome analysis of wood decay fungi *Postia placenta* and *Phanerochaete chrysosporium*. *Appl Environ Microbiol*. 2010;76(11):3599-3610.

78. Martinez D, Challacombe J, Morgenstern I, Hibbett D, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Duenas FJ, Martinez AT, Kersten P, et al. Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. *Proc Natl Acad Sci U S A*. 2009;106(6):1954-1959.

79. Foreman PK, Brown D, Dankmeyer L, Dean R, Diener S, Dunn-Coleman NS, Goedegebuur F, Houfek TD, England GJ, Kelley AS, et al. Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. *J Biol Chem*. 2003;278(34):31988-31997.

80. Harris PV, Welner D, McFarland KC, Re E, Navarro Poulsen JC, Brown K, Salbo R, Ding H, Vlasenko E, Merino S, et al. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry*. 2010;49(15):3305-3316.

81. Westereng B, Ishida T, Vaaje-Kolstad G, Wu M, Eijsink VGH, Igarashi K, Samejima M, Ståhlberg J, Horn SJ, Sandgren M. The putative endoglucanase *Pc*GH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. Uversky VN, ed. *PLoS One*. 2011;6(11):e27807.

82. Eibinger M, Ganner T, Bubner P, Rošker S, Kracher D, Haltrich D, Ludwig R, Plank H, Nidetzky B. Cellulose surface degradation by a lytic polysaccharide monooxygenase and its effect on cellulase hydrolytic efficiency. *J Biol Chem*. 2014;289(52):35929-35938.

83. Jung S, Song Y, Kim HM, Bae H. Enhanced lignocellulosic biomass hydrolysis by oxidative lytic polysaccharide monooxygenases (LPMOs) GH61 from *Gloeophyllum trabeum*. *Enzyme Microb Technol*. 2015;77:38-45.

84. Langston JA, Shaghasi T, Abbate E, Xu F, Vlasenko E, Sweeney MD. Oxidoreductive cellulose

depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl Environ Microbiol*. 2011;77(19):7007-7015.

85. Hu J, Chandra R, Arantes V, Gourlay K, van Dyk JS, Saddler JN. The addition of accessory enzymes enhances the hydrolytic performance of cellulase enzymes at high solid loadings. *Bioresour Technol*. 2015;186:149-153.
86. Harris PV, Xu F, Kreel NE, Kang C, Fukuyama S. New enzyme insights drive advances in commercial ethanol production. *Curr Opin Chem Biol*. 2014;19:162-170.
87. Hemsworth GR, Henrissat B, Davies GJ, Walton PH. Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol*. 2014;10(2):122-126.
88. Vu VV, Beeson WT, Span EA, Farquhar ER, Marletta MA. A family of starch-active polysaccharide monooxygenases. *Proc Natl Acad Sci*. 2014;111(38):13822-13827.
89. Hori C, Igarashi K, Katayama A, Samejima M. Effects of xylan and starch on secretome of the basidiomycete *Phanerochaete chrysosporium* grown on cellulose. *FEMS Microbiol Lett*. 2011;321(1):14-23.
90. Yakovlev I, Vaaje-Kolstad G, Hietala AM, Stefańczyk E, Solheim H, Fossdal CG. Substrate-specific transcription of the enigmatic GH61 family of the pathogenic white-rot fungus *Heterobasidion* irregulare during growth on lignocellulose. *Appl Microbiol Biotechnol*. 2012.
91. Poidevin L, Berrin JG, Bennati-Granier C, Levasseur A, Herpoël-Gimbert I, Chevret D, Coutinho PM, Henrissat B, Heiss-Blanquet S, Record E. Comparative analyses of *Podospora anserina* secretomes reveal a large array of lignocellulose-active enzymes. *Appl Microbiol Biotechnol*. 2014;98(17):7457-7469.
92. Berka RM, Grigoriev IV, Otillar R, Salamov A, Grimwood J, Reid I, Ishmael N, John T, Darmond C, Moisan MC, et al. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nat Biotechnol*. 2011;29(10):922-927.
93. Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martinez AT, Otillar R, Spatafora JW, Yadav JS, et al. The paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science*. 2012;336(6089):1715-1719.
94. Busk PK, Lange L. Classification of fungal and bacterial lytic polysaccharide monooxygenases. *BMC Genomics*. 2015;16(1):368.
95. Bentley SD, Chater KF, Cerdeño-Tárraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*. 2002;417(6885):141-147.

96. Paspaliari DK, Loose JSM, Larsen MH, Vaaje-Kolstad G. *Listeria monocytogenes* has a functional chitinolytic system and an active lytic polysaccharide monooxygenase. *FEBS J*. 2015;282(5):921-936.

97. Loose JSM, Forsberg Z, Fraaije MW, Eijsink VGH, Vaaje-Kolstad G. A rapid quantitative activity assay shows that the *Vibrio cholerae* colonization factor GbpA is an active lytic polysaccharide monooxygenase. *FEBS Lett*. 2014;588(18):3435-3440.

98. Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunæs AC, Stenstrøm Y, Mackenzie A, Sørlie M, Horn SJ, Eijsink VGH. Cleavage of cellulose by a CBM33 protein. *Protein Sci*. 2011;20(9):1479-1483.

99. Vaaje-Kolstad G, Bøhle LA, Gåseidnes S, Dalhus B, Bjørås M, Mathiesen G, Eijsink VGH. Characterization of the chitinolytic machinery of *Enterococcus faecalis* V583 and high-resolution structure of its oxidative CBM33 enzyme. *J Mol Biol*. 2012;416(2):239-254.

100. Dimarogona M, Topakas E, Christakopoulos P. Recalcitrant polysaccharide degradation by novel oxidative biocatalysts. *Appl Microbiol Biotechnol*. 2013;97(19):8455-8465.

101. Forsberg Z, Mackenzie AK, Sorlie M, Rohr ÅK, Helland R, Arvai AS, Vaaje-Kolstad G, Eijsink VGH. Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc Natl Acad Sci*. 2014;111(23):8446-8451.

102. Ghatge SS, Telke AA, Waghmode TR, Lee Y, Lee KW, Oh DB, Shin HD, Kim SW. Multifunctional cellulolytic auxiliary activity protein *Hc*AA10-2 from *Hahella chejuensis* enhances enzymatic hydrolysis of crystalline cellulose. *Appl Microbiol Biotechnol*. 2015;99(7):3041-3055.

103. Isaksen T, Westereng B, Aachmann FL, Agger JW, Kracher D, Kittl R, Ludwig R, Haltrich D, Eijsink VGH, Horn SJ. A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem*. 2014;289(5):2632-2642.

104. Bey M, Zhou S, Poidevin L, Henrissat B, Coutinho PM, Berrin JG, Sigoillot JC. Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monooxygenases (family GH61) from *Podospora anserina*. *Appl Environ Microbiol*. 2013;79(2):488-496.

105. Bennati-Granier C, Garajova S, Champion C, Grisel S, Haon M, Zhou S, Fanuel M, Ropartz D, Rogniaux H, Gimbert I, et al. Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by *Podospora anserina*. *Biotechnol Biofuels*. 2015;8(1):90.

106. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SWA, Koetsier MJ, van Berkel WJH,

Gruppen H, Kabel MA. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels*. 2015;8(1):101.

107. Lo Leggio L, Simmons TJ, Poulsen JCN, Frandsen KEH, Hemsworth GR, Stringer MA, von Freiesleben P, Tovborg M, Johansen KS, De Maria L, et al. Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nat Commun*. 2015;6:5961.
108. Li X, Beeson WT, Phillips CM, Marletta MA, Cate JHD. Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure*. 2012;20(6):1051-1061.
109. Wu M, Beckham GT, Larsson AM, Ishida T, Kim S, Payne CM, Himmel ME, Crowley MF, Horn SJ, Westereng B, et al. Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the basidiomycota fungus *Phanerochaete chrysosporium*. *J Biol Chem*. 2013;288(18):12828-12839.
110. Aachmann FL, Sorlie M, Skjak-Braek G, Eijsink VGH, Vaaje-Kolstad G. NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. *Proc Natl Acad Sci*. 2012;109(46):18779-18784.
111. Hemsworth GR, Taylor EJ, Kim RQ, Gregory RC, Lewis SJ, Turkenburg JP, Parkin A, Davies GJ, Walton PH. The copper active site of CBM33 polysaccharide oxygenases. *J Am Chem Soc*. 2013;135(16):6069-6077.
112. Beeson WT, Vu VV, Span EA, Phillips CM, Marletta MA. Cellulose degradation by polysaccharide monooxygenases. *Annu Rev Biochem*. 2015;84(1):923-946.
113. Hemsworth GR, Davies GJ, Walton PH. Recent insights into copper-containing lytic polysaccharide mono-oxygenases. *Curr Opin Struct Biol*. 2013;23(5):660-668.
114. Book AJ, Yennamalli RM, Takasuka TE, Currie CR, Phillips GN, Fox BG. Evolution of substrate specificity in bacterial AA10 lytic polysaccharide monooxygenases. *Biotechnol Biofuels*. 2014;7(1):109.
115. Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Várnai A, Røhr ÅK, Payne CM, Sørlie M, Sandgren M, et al. Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *J Biol Chem*. 2015;290(38):22955-22969.
116. Gudmundsson M, Kim S, Wu M, Ishida T, Momeni MH, Vaaje-Kolstad G, Lundberg D, Royant A, Stahlberg J, Eijsink VGH, et al. Structural and electronic snapshots during the transition from a Cu(II) to Cu(I) metal center of a lytic polysaccharide monooxygenase by X-ray photo-reduction. *J Biol Chem*. 2014;289(27):18782-18792.

117. Hervé C, Rogowski A, Blake AW, Marcus SE, Gilbert HJ, Knox JP. Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *Proc Natl Acad Sci*. 2010;107(34):15293-15298.

118. Carrard G, Koivula A, Söderlund H, Béguin P. Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. *Proc Natl Acad Sci U S A*. 2000;97(19):10342-10347.

119. Kim S, Stahlberg J, Sandgren M, Paton RS, Beckham GT. Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. *Proc Natl Acad Sci*. 2014;111(1):149-154.

120. Beeson WT, Phillips CM, Cate JHD, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc*. 2012;134(2):890-892.

121. Solomon EI, Heppner DE, Johnston EM, Ginsbach JW, Cirera J, Qayyum M, Kieber-Emmons MT, Kjaergaard CH, Hadt RG, Tian L. Copper active sites in biology. *Chem Rev*. 2014;114(7):3659-3853.

122. Kjaergaard CH, Qayyum MF, Wong SD, Xu F, Hemsworth GR, Walton DJ, Young NA, Davies GJ, Walton PH, Johansen KS, et al. Spectroscopic and computational insight into the activation of O_2 by the mononuclear Cu center in polysaccharide monooxygenases. *Proc Natl Acad Sci U S A*. June 2014.

123. Courtade G, Balzer S, Forsberg Z, Vaaje-Kolstad G, Eijsink VGH, Aachmann FL. ^{1}H, ^{13}C, ^{15}N resonance assignment of the chitin-active lytic polysaccharide monooxygenase *Bl*LPMO10A from *Bacillus licheniformis*. *Biomol NMR Assign*. 2015;9(1):207-210.

124. Forsberg Z, Røhr AK, Mekasha S, Andersson KK, Eijsink VGH, Vaaje-Kolstad G, Sørlie M. Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases. *Biochemistry*. 2014;53(10):1647-1656.

125. Forsberg Z, Nelson CE, Dalhus B, Mekasha S, Loose JSM, Crouch LI, Røhr ÅK, Gardner JG, Eijsink VGH, Vaaje-Kolstad G. Structural and functional analysis of a lytic polysaccharide monooxygenase important for efficient utilization of chitin in *Cellvibrio japonicus*. *J Biol Chem*. 2016;2010(1):jbc.M115.700161.

126. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels*. 2012;5(1):79.

127. Vu VV, Beeson WT, Phillips CM, Cate JHD, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc*. 2014;136(2):562-565.

128. Sygmund C, Kracher D, Scheiblbrandner S, Zahma K, Felice AKG, Harreither W, Kittl R, Ludwig R. Characterization of the two *Neurospora crassa* cellobiose dehydrogenases and their connection to oxidative cellulose degradation. *Appl Environ Microbiol*. 2012;78(17):6161-6171.

129. Kracher D, Zahma K, Schulz C, Sygmund C, Gorton L, Ludwig R. Inter-domain electron transfer in cellobiose dehydrogenase: modulation by pH and divalent cations. *FEBS J*. 2015;282(16):3136-3148.

130. Gardner JG, Crouch L, Labourel A, Forsberg Z, Bukhman Y V., Vaaje-Kolstad G, Gilbert HJ, Keating DH. Systems biology defines the biological significance of redox-active proteins during cellulose degradation in an aerobic bacterium. *Mol Microbiol*. 2014;94(5):1121-1133.

131. Dimarogona M, Topakas E, Olsson L, Christakopoulos P. Lignin boosts the cellulase performance of a GH-61 enzyme from *Sporotrichum thermophile*. *Bioresour Technol*. 2012;110:480-487.

132. Hu J, Arantes V, Pribowo A. Substrate factors that influence the synergistic interaction of AA9 and cellulases during the enzymatic hydrolysis of biomass. *Energy Environ Sci*. 2014;7:2308-2315.

133. Gougoulias C, Clark JM, Shaw LJ. The role of soil microbes in the global carbon cycle: Tracking the below-ground microbial processing of plant-derived carbon for manipulating carbon dynamics in agricultural systems. *J Sci Food Agric*. 2014;94(12):2362-2371.

134. Mäkelä MR, Donofrio N, de Vries RP. Plant biomass degradation by fungi. *Fungal Genet Biol*. 2014;72:2-9.

135. Glass NL, Schmoll M, Cate JHD, Coradetti S. Plant cell wall deconstruction by ascomycete fungi. *Annu Rev Microbiol*. 2013;67:477-498.

136. Cragg SM, Beckham GT, Bruce NC, Bugg TD, Distel DL, Dupree P, Etxabe AG, Goodell BS, Jellison J, McGeehan JE, et al. Lignocellulose degradation mechanisms across the Tree of Life. *Curr Opin Chem Biol*. 2015;29:108-119.

137. de Vries RP, Visser J. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. *Microbiol Mol Biol Rev*. 2001;65(4):497-522.

138. Navarro D, Rosso MN, Haon M, Olivé C, Bonnin E, Lesage-Meessen L, Chevret D, Coutinho PM, Henrissat B, Berrin JG. Fast solubilization of recalcitrant cellulosic biomass by the basidiomycete fungus *Laetisaria arvalisinvolves* successive secretion of oxidative and hydrolytic enzymes. *Biotechnol Biofuels*. 2014;7(1):143.

139. Østergaard LH, Olsen HS. Industrial applications of fungal enzymes. In: *The Mycota*.;

2011:269-290.

140. Sánchez C. Lignocellulosic residues: Biodegradation and bioconversion by fungi. *Biotechnol Adv*. 2009;27(2):185-194.

141. Grigoriev IV, Cullen D, Goodwin SB, Hibbett D, Jeffries TW, Kubicek CP, Kuske C, Magnuson JK, Martin F, Spatafora JW, et al. Fueling the future with fungal genomics. *Mycol An Int J fungal Biol*. 2011;2(3):192-209.

142. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, et al. MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Res*. 2014;42(D1):699-704.

143. Bianco L, Perrotta G. Methodologies and perspectives of proteomics applied to filamentous fungi: From sample preparation to secretome analysis. *Int J Mol Sci*. 2015;16(3):5803-5829.

144. Caccia D, Dugo M, Callari M, Bongarzone I. Bioinformatics tools for secretome analysis. *Biochim Biophys Acta - Proteins Proteomics*. 2013;1834(11):2442-2453.

145. Sperschneider J, Williams AH, Hane JK, Singh KB, Taylor JM. Evaluation of secretion prediction highlights fiffering approaches needed for oomycete and fungal effectors. *Front Plant Sci*. 2015;6(1168):1-14.

146. Choi J, Park J, Kim D, Jung K, Kang S, Lee YH. Fungal secretome database: Integrated platform for annotation of fungal secretomes. *BMC Genomics*. 2010;11(105):1-15.

147. Hori C, Gaskell J, Igarashi K, Kersten P, Mozuch M, Samejima M, Cullen D. Temporal alterations in the secretome of the selective ligninolytic fungus *Ceriporiopsis subvermispora* during growth on aspen wood reveal this organism's strategy for degrading lignocellulose. *Appl Environ Microbiol*. 2014;80(7):2062-2070.

148. Liu D, Li J, Zhao S, Zhang R, Wang M, Miao Y, Shen Y, Shen Q. Secretome diversity and quantitative analysis of cellulolytic *Aspergillus fumigatus* Z5 in the presence of different carbon sources. *Biotechnol Biofuels*. 2013;6(1):149.

149. Saykhedkar S, Ray A, Ayoubi-Canaan P, Hartson SD, Prade R, Mort AJ. A time course analysis of the extracellular proteome of *Aspergillus nidulans* growing on sorghum stover. *Biotechnol Biofuels*. 2012;5(1):52.

150. Phillips CM, Iavarone AT, Marletta MA. Quantitative proteomic approach for cellulose degradation by *Neurospora crassa*. *J Proteome Res*. 2011;10(9):4177-4185.

151. Latgé JP. The pathobiology of Aspergillus fumigatus. *Trends Microbiol*. 2001;9(8):382-389.

152. Baker SE. *Aspergillus niger* genomics: Past, present and into the future. *Med Mycol*. 2006;44:S17-S21.

153. Galagan JE, Calvo SE, Cuomo C, Ma LJ, Wortman JR, Batzoglou S, Lee SI, Baştürkmen M, Spevak CC, Clutterbuck J, et al. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature*. 2005;438(7071):1105-1115.

154. Coutinho PM, Andersen MR, Kolenova K, VanKuyk PA, Benoit I, Gruben BS, Trejo-Aguilar B, Visser H, van Solingen P, Pakula T, et al. Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. *Fungal Genet Biol*. 2009;46:S161-S169.

155. Wortman JR, Gilsenan JM, Joardar V, Deegan J, Clutterbuck J, Andersen MR, Archer D, Bencina M, Braus G, Coutinho P, et al. The 2008 update of the *Aspergillus nidulans* genome annotation: a community effort. *Fungal Genet Biol*. 2009;46(1):S2-S13.

156. Arnaud MB, Cerqueira GC, Inglis DO, Skrzypek MS, Binkley J, Chibucos MC, Crabtree J, Howarth C, Orvis J, Shah P, et al. The *Aspergillus* Genome Database (AspGD): Recent developments in comprehensive multispecies curation, comparative genomics and community resources. *Nucleic Acids Res*. 2012;40(D):653-659.

157. Schneider T, Gerrits B, Gassmann R, Schmid E, Gessner MO, Richter A, Battin T, Eberl L, Riedel K. Proteome analysis of fungal and bacterial involvement in leaf litter decomposition. *Proteomics*. 2010;10(9):1819-1830.

158. Couturier M, Navarro D, Olivé C, Chevret D, Haon M, Favel A, Lesage-Meessen L, Henrissat B, Coutinho PM, Berrin JG. Post-genomic analyses of fungal lignocellulosic biomass degradation reveal the unexpected potential of the plant pathogen *Ustilago maydis*. *BMC Genomics*. 2012;13(1):57.

159. Nelson PE, Dignani MC, Anaissie EJ. Taxonomy, biology, and clinical aspects of *Fusarium* species. *Clin Microbiol Rev*. 1994;7(4):479-504.

160. Goswami RS, Kistler HC. Heading for disaster: *Fusarium graminearum* on cereal crops. *Mol Plant Pathol*. 2004;5(6):515-525.

161. Dean R, Van Kan JAL, Pretorius ZA, Hammond-Kosack KE, Di Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J, et al. The Top 10 fungal pathogens in molecular plant pathology. *Mol Plant Pathol*. 2012;13(4):414-430.

162. King R, Urban M, Hammond-Kosack MCU, Hassani-Pak K, Hammond-Kosack KE. The completed genome sequence of the pathogenic ascomycete fungus *Fusarium graminearum*. *BMC Genomics*. 2015;16(1):544.

163. Zhao Z, Liu H, Wang C, Xu J. Correction: Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics*. 2014;15(1):6.

164. Paper JM, Scott-Craig JS, Adhikari ND, Cuomo CA, Walton JD. Comparative proteomics of extracellular proteins *in vitro* and *in planta* from the pathogenic fungus *Fusarium graminearum*. *Proteomics*. 2007;7(17):3171-3183.

165. Debeire P, Delalande F, Habrylo O, Jeltsch JM, Van Dorsselaer A, Phalip V. Enzymatic cocktails produced by *Fusarium graminearum* under submerged fermentation using different lignocellulosic biomasses. *FEMS Microbiol Lett*. 2014;355(2):116-123.

CHAPTER 2. PAPER I

Lytic polysaccharide monooxygenases and other oxidative enzymes are abundantly secreted by *Aspergillus nidulans* grown on different starches

Lytic polysaccharide monooxygenases and other oxidative enzymes are abundantly secreted by *Aspergillus nidulans* grown on different starches

Laura Nekiunaite[1a], Magnus Ø. Arntzen[2a], Birte Svensson[1], Gustav Vaaje-Kolstad[2*] and Maher Abou Hachem[1*]

1 Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark
2 Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Aas, Norway
a These authors contributed equally to this work
* Correspondence: maha@bio.dtu.dk; gustav.vaaje-kolstad@nmbu.no

Abstract

Background:

Starch is the second most abundant renewable terrestrial plant-derived biomass and a major feedstock in non-food industrial applications and first generation biofuel production. In contrast to lignocellulose biomass, detailed insight into fungal degradation of starch is currently lacking. This study explores the secretomes of *Aspergillus nidulans* grown on cereal starches from wheat and high amylose (HA) maize, and legume starch from pea for five days.

Results:

A. nidulans grew efficiently on cereal starches, whereas growth on pea starch was poor. The secretomes at days 3–5 were dynamic, but starch-type dependent, as also reflected by amylolytic activity measurements. Nearly half of the 312 proteins in the secretomes were carbohydrate active enzymes (CAZymes), most of which glycoside hydrolases (GHs) and oxidative auxiliary activities (AAs). The secretomes from maize starch contained the highest number of proteins and diversity of CAZymes. By contrast, pea starch secretomes were relatively richer in proteases and uncharacterized enzymes. The abundance of the α-amylase (AmyB) decreased with time, whereas an increase was noted for other starch degrading enzymes, *e.g.* AmyF and the glucoamylase (GlcA) in the last three days of culture. Two AA13 LPMOs, co-secreted with amylolytic enzymes, were

amongst the most abundant CAZymes in the culture supernatants. One of these two enzymes possessed a starch binding module of CBM20 and was the most abundant protein in the starch-binding fraction of the secretome. Binding sites of the AmyR transcriptional-regulator of amylolytic enzymes in *A. nidulans* were identified upstream both AA13 genes. Several redox-active enzymes were also co-secreted abundantly including AA9 LPMOs, as well as AA3, AA7, and other non-CAZyme oxidoreductases.

Conclusions:

Co-secretion and the high abundance of AA13 LPMOs are indicative of a key role in starch deconstruction. The increase in AA13 LPMOs abundance with time may reflect the accumulation of the more resistant starch fraction towards the later stages of the culture. The identification of AmyR sites upstream AA13 LPMOs unveils a co-regulation mechanism of LPMOs featuring in starch utilization. This study demonstrates for the first time the biological significance of LPMOs in starch degradation and shows a direct mechanism for co-regulation of LPMOs and GHs in this process.

Keywords: *Aspergillus nidulans*, Biofuels, Carbohydrate-active enzymes, Carbohydrate binding module family 20 (CBM20), Filamentous fungi, Lytic polysaccharide monooxygenase (LPMO), Starch, Starch-binding.

BACKGROUND

Starch is one of the most abundant renewable biopolymers in nature [1]. Annually two billion tons of starch crops are harvested worldwide, making it an attractive resource for industrial applications such as production of first generation biofuels, pharmaceuticals, textiles, detergents, paper, and food [1]. Starch consists of the two α-glucan polymers, amylose and amylopectin. Amylose is an essentially linear polymer of α-(1,4)-linked glucosyl units, while amylopectin which constitutes 70−80% of starch granules, is a branched macromolecule having α-(1,4)-glucan chains branched with approximately 5% α-(1,6)-glucosidic linkages [2]. Despite this chemical simplicity, the α-glucan chains are arranged radially in a supramolecular assembly forming water insoluble granules varying in size (1−150 μm), morphology, crystal-type packing and crystallinity (15−45%) [3,4]. This organization renders starch, especially its crystalline regions, into a relatively challenging substrate for complete enzymatic deconstruction. The extent of starch resistance to enzymatic hydrolysis is correlated with the botanical origin and processing, both factors having influence on crystal packing, crystallinity and supramolecular structure of the starch granule.

The high utilization potential of starch as a renewable biological resource and an industrial feedstock motivates efforts to improve starch hydrolysis yields, particularly from more resistant starch types and for shortening process times. Gain in yields of hydrolysis would have a considerable impact on efficiency and cost of industrial starch processing as well as the reduction of environmental impact of this process [5].

The classical paradigm of polysaccharide degradation by hydrolytic enzymes, which has been valid for decades, was recently re-visited by the discovery of the copper-dependent lytic polysaccharide monooxygenases (LPMOs) [6,7]. LPMOs catalyze the oxidative cleavage of glycosidic linkages of polysaccharides in the presence of molecular oxygen and an external electron donor, by hydroxylation of either the C1 or C4 carbon of the glycosidic bond [8,9]. These enzymes play an instrumental role in degradation of recalcitrant crystalline polysaccharides such as cellulose [7,10]

and chitin [6], which renders LPMOs into pivotal tools in industrial biomass conversion [11]. Recently it has been established that some LPMOs oxidize non-crystalline hemicelluloses and soluble cello-oligosaccharides [12–14]. Moreover, the recent discovery of starch-active LPMOs [15,16], which have been assigned into auxiliary activity family 13 (AA13) in the CAZy database [17], suggests that these enzymes play a role in starch degradation together with amylolytic hydrolases.

Filamentous fungi produce impressive amounts of hydrolytic enzymes targeting polysaccharides. In addition to glycoside hydrolyses (GHs), several fungi also produce a multitude of LPMOs [18,19]. In contrast to lignocellulose matrices, little is known regarding the type and composition of enzyme cocktails deployed by fungi for the degradation of starches differing in botanical origin and properties. Notably, the involvement of oxidative enzymes including LPMOs in starch degradation has not been demonstrated *in vivo*. Here, we provide new insight into the excellent starch-degrading capabilities of the well-studied, genome sequenced saprophytic ascomycete *Aspergillus nidulans*, that is taxonomically related to well established industrial cell factory species such as *Aspergillus niger and Aspergillus oryzae* [20]. By integrating secretomics and enzyme activity assays, we analyzed temporal changes of the enzymes secreted by *A. nidulans* to sustain growth on three different starches in the course of 5 days. The data demonstrate differences in growth on the different starch substrates, which is also reflected in the enzymes secreted by the fungus in these cultures. A common feature of growth on starch was that two AA13 LPMOs including the modular starch-specific enzyme joint to a starch binding domain of family 20 (CBM20) were amongst the top most abundant enzymes together with a variety of LPMOs and other oxidative enzymes. This finding suggests that oxidative cleavage of α-glucosidic bonds plays a significant role in starch breakdown. Altogether, the novel insight into enzymatic activities secreted by *A. nidulans* and related fungi for efficient starch breakdown is relevant for design of enzyme mixtures with enhanced bioconversion efficiencies of starches especially those resistant to hydrolytic degradation.

RESULTS

Starch substrates and fungal growth

To assess the ability of *A. nidulans* to sense differences in the origin and structure of the starch substrates and to fine-tune the composition of secreted enzymes accordingly, this fungus was grown on wheat, high-amylose (HA) maize and pea starches, and the resulting secretomes were analyzed.

A. nidulans grew efficiently on wheat and HA maize starch and no intact starch granules were distinguished from the fungal biomass after 5 days, suggesting extensive degradation of both starches. By contrast, growth was poor on pea starch, leaving significant amounts of intact residual starch granules at culture harvest, which clearly demonstrates important differences due to the botanical origin and properties of the starch on enzymatic deconstruction and growth.

Enzymatic analysis of amylolytic activities

The α-amylase and α-glucosidase activities were measured in the filtered culture supernatants. The average activities of the biological triplicates at day 1 to 5 in different starch media are shown in Fig. 1. Both the α-amylase and α-glucosidase activities in the supernatants were substrate dependent. The highest α-amylase and α-glucosidase activities were measured in the wheat and maize starch culture supernatants, respectively. The α-amylase activity increased to a maximum in 3–4 days and decreased thereafter, with activity maximum (0.21 U/ml) after 4 days in wheat starch (Fig. 1A). By contrast, the α-amylase activity in the pea starch culture supernatants was barely detectable, consistent with the poor growth on this substrate.

The α-glucosidase activity increased over time in all samples and showed the highest activity (53 U/ml) in HA maize starch medium on day 5 (Fig. 1B). The glucosidase activity was roughly similar between day 1–3 in the pea starch culture, whereas a one-fold increase was observed at days 4–5. By contrast, the α-glucosidase activities in the HA maize and wheat cultures increased steadily until day 5 (Fig. 1B).

Survey of secreted *A. nidulans* proteins

Filtered supernatants from *A. nidulans* cultures grown on wheat, HA maize and pea starch media were analyzed using liquid chromatography combined with tandem mass spectrometry (LC-MS/MS). The analysis of the data set (Additional file 1: Table S1) revealed dynamic secreted protein profiles of *A. nidulans* grown on these different starches over the course of 5 days. The complete proteome of *A. nidulans* contains 10556 sequences of which 9.7% are predicted to be secreted when using a combination of three different prediction algorithms. In this study, 937 proteins were identified, of which 33% are predicted to be secreted, which reflects a clear enrichment of secreted proteins compared to the complete proteome. The secreted proteins on days 3, 4 and 5 were identified and assigned to different functional categories including proteases and various carbohydrate active proteins and clustered both according to abundance and trend related to increase/decrease over time (Additional file 3: Figures S1 and S2, respectively). The number of secreted proteins detected in each culture supernatant varied between 174 (pea starch, day 5) and 221 (HA maize starch, day 1) and generally the number of identified proteins decreased at day 5, as compared to days 3 and 4, but less so in pea (≈ 4.4%) followed by wheat (≈ 7.5%) and maize (≈ 9%) starches (Fig. 2). Approximately 20% of the secreted proteins were assigned as uncharacterized based on the lack of characterized homologous counterparts. For the remaining secretome, CAZymes (carbohydrate active enzymes and proteins assigned into the CAZy database, http://www.cazy.org) represented the largest category amounting to 44% of the secretome (Fig. 3).

Approximately 30% of the 478 annotated CAZymes were identified in the secretomes despite growth on purified starches and not crude plant material. This data revealed that most of the identified CAZymes are GHs or AAs. The proportion of the different CAZyme modules was fairly similar between the different starches, but generally the total number of identified CAZymes was lower in the pea starch cultures (Figs. 2 and 3). *A. nidulans* secreted multiple enzymes (4–10

enzymes) from certain CAZy families *e.g*. GH3, GH16 and GH43, and CE10 as well as AA3, AA7 and AA9, with the three latter families harboring oxidative activities (Additional file 2: Table S2). Some ancillary carbohydrate module (CBM) families were identified including four CBM20s which mediate starch binding in addition to cellulose-specific (CBM1 and CBM6), different β-glucans (CBM6, CBM24 and CBM43), chitin (CBM18), and arabinofuranosyl (CBM42) binding modules.

Cell wall degrading activities from GH5, GH6, GH7, GH53 and GH62 were present only in wheat and HA maize cultures. Beyond this, only a few families were not identified in all samples, including GH35 which was only observed in HA maize and pea, and GH24 and GH95 that only were identified in HA maize and pea induced samples, respectively (Additional file 2: Table S2). Notably, amongst the putative carbohydrate active enzymes, a GH13 α-amylase (AmyF, Q5B7U2), a GH31 α-glucosidase (AgdB, G5EB11) and an AA13 LPMO (*An*LPMO13B, Q5B027) were the most abundant in all samples. The analysis revealed the secretion of six different putative LPMOs belonging to families AA9 and AA13 in all starch samples, and one AA11, which was only found in the pea starch culture on day 3 (Table 1, Additional file 2: Table S2).

Starch degrading enzymes

The GH families 13, 15 and 31 contain α-amylases (and other α-glucan acting enzymes), glucoamylases and α-glucosidases, respectively. Seven of the twelve predicted extracellular enzymes were assigned into these families and identified in the secretome (Table 2, Fig. 4). Three putative GH13 α-amylases (G5EB45, G5EAT0, and Q5B7U2), and three GH31 α-1,4-glucosidases (G5EB03, G5EB11, and Q5BET9) were present in all samples on all days. A GH15 α-glucoamylase (C8VLL3) was identified in all wheat and HA maize starch medium and only on day 4 and day 5 in pea. Moreover, the recently functionally characterized starch degrading AA13 LPMO [16] was identified in addition to a second AA13 enzyme (Table 1, Fig. 4). Interestingly, the GH13 α-amylase having a CBM20 (G5EAT0) shows a declining tendency over time (Fig. 4, Additional file 3: Figure S2).

Otherwise, all amylolytic proteins are highly abundant in all time points or show an increase over time (Fig. 4, Additional file 3: Figure S2).

Redox-active enzymes

The *A. nidulans* genome contains 13 genes encoding LPMOs: 9 AA9s, 2 AA11s and 2 AA13s (Table 1). Two AA9 LPMOs are assigned in clade LPMO1, one in LPMO2, five in LPMO3 clade and one in an unclassified cluster, as phylogenetically defined by Book et al. (Table 1) [21]. None of the five clade 3 LPMOs were identified in the *A. nidulans* supernatants (Table 1), whereas the remaining four were all observed, albeit with low abundance for enzymes (Fig. 3). The two AA13 LPMOs were amongst the most abundant proteins in the culture supernatants, *An*LPMO13B seeming at higher abundance than *An*LPMO13A (Q5B1W7). Notably, the abundance of both proteins increased slightly over time (Fig. 4, Additional file 3: Figure S2).

In addition to LPMOs, the secretomes also contain a variety of other redox active enzymes, including a catalase, laccase and thioredoxin reductase as well as members of AA3 and AA7 (Fig. 4, Additional file 3: Figure S2). The 6 AA3s secreted by *A. nidulans* (out of 16) belong to sub group 2 that accommodates both aryl alcohol and glucose-1 oxidases. Of these, two (Q5AV48 and Q5ALN2) were constantly highly abundant, whereas C8VDT4 and C5AUN2 showed a substantial increase in abundance over time (Fig. 4, Additional file 3: Figure S2). The *A. nidulans* secretomes show the presence of 10 AA7 putative glucooligosaccharide oxidases (GOOs), three of which (Q5AY23, Q5B7X9 and C8VCU1) were highly abundant in most samples, the two former and the latter showing a decreasing and increasing trend over time, respectively (Fig. 4).

The secretomes contain several non-CAZyme redox-active enzymes. As a matter of fact, the most abundant protein in all collected supernatants (except day 5 maize and day 3 pea) is Catalase B (CatB) (P78619), which also shows a slight decrease in abundance over time in wheat and maize, and an opposite trend in pea (Fig. 3, Additional file 2: Tables S3–S5). Thioredoxin reductase is

abundant in most samples (except cultures grown on maize starch) and shows a decreasing trend in the wheat and maize samples like CatB. A laccase (Q5BEX9) is also present in the secretome at low abundance, and more so over time.

Plant cell wall degrading non-oxidative enzymes

Ten putative β-glucoside and cellulose-degrading enzymes were identified. A GH1 (Q5AR97) and five putative GH3 β-glucosidases (Q5B5S8, Q5B6C6, Q5AWD4, Q5B7X0 and Q5B9F2) were identified in all samples. Two of these enzymes (Q5AR97, Q5B7X0) appeared only after 4 days in the pea starch. Notably, cellobiohydrolases (GH6) and endoglucanases (GH5, 7) active on cellulose were not identified in the pea starch culture supernatant. Two GH5 (Q5AUG2, Q5BDU5) were only present in the maize, whereas the GH6 (C8VSG6) and GH7 (Q5B7R2) were also identified in wheat starch cultures. Twelve putative hemicellulose degrading enzymes were identified including a GH10 (Q00177) and a GH11 (P55332) β-1,4-endoxylanases and four putative GH43 β-xylosidases/α-L-arabinofuranosidases (C8VCT5, C8VKG9, Q5AUM3 and Q5AV99). α-Arabinofuranosidases of GH54 (O74288) and GH62 (Q5AUX2) were also identified, the latter only in wheat and HA maize starch media together with a GH53 β-1,4-endogalactanase (Q5B153). Moreover, a putative GH93 exo-arabinanase (Q5BBM0), a GH27 (Q5AX28) and a GH36 (Q5AU92) α-1,6-galactosidases were identified. Putative pectin degrading enzymes including a GH2 β-glucuronidase (Q5BAN5), a GH105 unsaturated rhamnogalacturonan hydrolase (Q5AQP7), and two putative GH28 exo-polygalacturonases (Q5ASG9 and Q873X6) were detected, the latter only in HA maize starch cultures. Finally, a GH35 putative β-galactosidase (Q5BFC4) was detected in HA maize and pea starch cultures.

Fungal cell-wall active enzymes

Fifteen secreted enzymes putatively involved in fungal cell wall modification and degradation were identified including six GH16 glycanases (C8VUN8, Q5AY11, Q5AYL0, Q5BAP5, Q5BGT5 and Q5B4L5), a GH17 β-1,3-endoglucanase (Q5AUT0), two putative GH55 β-1,3-exoglucanase (Q5B3Q5

and C8VQV2), three GH72 β-1,3-glucanosyltransferases (C8VSK8, Q5AVM3 and Q5AW19), a GH81 endo-β-1,3-endoglucanase (C8VT57) and a putative GH20 *N*-acetylglucosaminidase (G5EB27). The putative GH71 α-1,3-glucanase (C8VKQ6) was detected only on days 4 and 5 in all cultures. In addition, three α-mannan degrading enzymes, one GH47 (Q5BF93) and two GH92 (C8VSU5 and Q5BAV5) were identified.

Purification of starch binding enzymes from wheat starch cultures

β-Cyclodextrin affinity chromatography was used to capture the fraction of *A. nidulans* secreted enzymes that possess affinity for starch. The elution fraction from this purification was analyzed by SDS-PAGE (Fig. 5) and dominant protein bands were identified using MALDI-TOF mass spectrometry (Table 3). The common feature to all proteins identified on the gel is the presence of CBM20 [22]. The three most intense bands contain the GH15 α-glucoamylase (AN7402), the GH13 subfamily1 α-amylase (G5EAT0), and the AA13 LPMO (Q5B1W7), that appeared as the most abundant protein in this starch-associated secretome fraction. The CBM20 of the α-amylase was also identified in two lower molecular mass bands on the gel, which is likely due to proteolytic cleavage (Fig. 5).

DISCUSSION

The abundance of starch from terrestrial biomass is only surpassed by cellulose. An excess of two billion tons of starch are annually produced from cereals and coarse grains and another 700 million tons from roots and tubers (The food and agriculture organization of the united nations [FAO]; http://www.fao.org/). The bulk of extracted starch from these crops is used for non-food industrial applications and as a feedstock for first generation biofuel that still contributes more than two thirds of the total biofuel production [23]. Technical and economic issues continue to hamper a larger transition to second generation lignocellulose based feedstocks [24], which together with the large investments made in first generation production facilities is likely to maintain the demand for starch as a feedstock in the time to come. Therefore, improving the efficiency of starch

depolymerization for both biofuel production and in other industrial applications that rely on starch degradation has significant economic and environmental potential.

The synergistic action of LPMOs with hydrolases [10,25] in the breakdown of recalcitrant polysaccharides has spurred interest in the exploitation of these enzymes in biofuel production [11]. Several proteomics studies revealed that LPMOs are highly co-secreted with hydrolases during fungal growth on plant derived biomass or purified polysaccharides [18,26]. The oxidative activity of starch specific LPMOs from *Neurospora crassa* [15] and *A. nidulans* [16] towards resistant retrograded starch was recently demonstrated and the presence of the *A. nidulans* LPMO elicited a 100-fold increase in the release of maltose by β-amylase on the same substrate [16]. This *in vitro* activity suggests that starch active AA13 LPMOs may play an important role in starch depolymerization by distinct fungi. Currently, there is no data on the involvement and biological significance of LPMOs in starch degradation. In this study we investigate the protein inventory secreted by *A. nidulans* during growth on three different starches differing in botanical origin and properties and show that starch specific LPMOs are highly abundant in all starch secretomes together with LPMOs of different specificities and other oxidative enzymes.

A. nidulans secretomes are fine-tuned to the different starch matrices

Fungal growth was far better on cereal starches from wheat and maize, while the pea starch was a poor growth substrate. The overall secretion profiles differed between the three substrates, but were more similar between the botanically closer cereal starches as compared to pea (Fig. 3, Additional file 3: Figure S1). This was also supported by enzyme assays that revealed generally much higher α-glucosidase and α-amylase activities in the cereal starch cultures and a substantial increase of α-amylase activity at day 3 that was not observed in the pea starch culture (Fig. 1). This is also consistent with large differences in the 20 most abundant proteins (Additional file 2: Tables S3–S5). The diversity of identified CAZymes was the highest in HA maize starch and the lowest in pea starch secretomes that contained a larger relative proportion of unknown enzymes (Fig. 2). These data

highlight the sensitivity of the regulatory machinery of the fungus and the fine-tuning of the secretomes to each starch substrate. The three used starches are different in their fine structures and composition. The HA maize starch has the highest amylose content with approximately 70% (w/w), followed by pea starch (>45% w/w) and wheat (25% w/w). HA maize starch granules are the smallest (5–25 µm), followed by pea (5–45 µm) and wheat (20–35 µm) [3,4]. Given the chemical simplicity of starch, it is unlikely that the secretomes merely reflect the differences in starch structure *per se*, but also the presence of other non-starch components including *e.g*. polysaccharides, lipids, proteins, which may impede the action of amylolytic enzymes [27]. The starch granules and protein storage vacuoles in monocot cereals like wheat and maize reside in the endosperm, while in pea the cotyledon tissue serves the same storage function [1,28]. Pea starch isolation has been reported to be more difficult due to the presence of fine fiber and insoluble flocculent proteins, significant amounts of which remain associated with the starch [29]. Indeed, numerous proteases are identified in the secretomes and more so in the pea than the cereal starch cultures (Fig. 2, Additional file 3: Figure S1). Pea starch is generally less well studied with respect to microbial degradation and the reasons for the poorer growth on this substrate are unclear. Starch from HA maize has been shown to be the least efficiently degraded by an α-amylase from *Aspergillus fumigatus*, followed by starch from pea [30]. The excellent growth of *A. nidulans* on the more resistant HA maize starch supports the possibility that non-starch components in the pea starch matrix rather than the recalcitrance of the starch is responsible for the poor growth. Indeed, legume cotyledons including pea are known to harbor a range of protease and hydrolase inhibitors [31], which may represent more challenging growth conditions and delay the deconstruction of the starch.

The major components in the cereal endosperm cell walls are typically arabinoxylans and mixed linkage glucans, in addition to small amounts of cellulose, xyloglucan and pectins [32]. By contrast, xyloglucan is the main structural polysaccharide in cotyledon cell walls [33]. This is consistent with

the presence of cellulases of GH5, GH6, GH7, in addition to GH62 arabinoxylan arabinofuranosidase exclusively in the cereal starch samples (Fig. 3, Additional file 2: Table S2). Xylose has been suggested to be the inducer of several polysaccharide degrading enzymes in aspergilli [34]. Possibly the co-secretion of a range of hemicellulose degrading enzymes including a GH10 xylanase with other amylolytic enzymes (Fig. 3, Additional file 3: Figure S2) is suggestive of the presence of xylose derived either from arabinoxylan in cereals or xyloglucan remaining in the legume pea starch preparations. This co-secretion may contribute to the degradation of cell-wall components and render starch more accessible to the action of amylolytic enzymes. An "anti-caging" effect of xylanases and other cell wall hydrolases has been observed *in vitro* starch digestion studies [35,36]. Another observation that supports the presence of cell wall components in the starch matrices is the secretion of five LPMOs of AA9 and 11 (including an enzyme with a putative cellulose binding module, Table 1), which feature in the breakdown of cellulose and hemicellulose [7,13,14]. These findings motivate further studies on the effect of low concentrations of key hemicellulose polysaccharide degrading enzymes on the yields of starch degradation.

The hydrolytic starch degrading machinery

The genomic and enzymatic capabilities of *A. nidulans* with regard to complex polysaccharide degradation [37,38] manifest the saprophytic lifestyle of this taxon. The secretomes of this fungus have also been investigated on lignocellulosic substrates [19]. The starch degradation model in fungi has been based on the hydrolytic biochemically well described amylolytic hydrolases from aspergilli [30,39]. The present study, however, is the first to report high resolution secretome data of the model fungus *A. nidulans* in response to growth on starches and to investigate the impact of starch origin on the secretomes.

The data revealed the secretion of a core of starch hydrolytic battery of *A. nidulans* comprising three α-amylases, three α-glucosidases, and a glucoamylase (Table 2). The absence of AmyD and AmyE of GH13_1 from the secretome (Table 2) is in excellent agreement with the role of these

enzymes in fungal cell wall α-glucan remodeling [40]. An intriguing observation is the decrease in abundance of the GH13-CBM20 (AmyB, G5EAT0) over time (Fig. 4, Additional file 3: Figure S2). The capture of the CBM20 proteolytic degradation product of this enzyme from the starch-binding fraction of the secretome (Fig. 5) proves proteolytic degradation has occurred, which is also in agreement with the decrease of α-amylase activity on day 5 in cereal starches, despites the slightly increased abundance of the α-amylase (AmyF) that lacks a CBM20. The pattern of increased abundance is also observed for GH15 glucoamylase (GlaA) and the CBM20 containing LPMO (*An*LPMO13A, Q5B1W7) (Fig. 4, Additional file 3: Figure S2). The different abundance patterns may suggest that different amylolytic activities are deployed depending on the availability of different substrate structures during the degradation of starch, but further work is needed to understand the significance of these observations.

Abundant deployment of LPMOs in starch degradation by *A. nidulans*

Several studies investigating the secretomes of plant biomass degrading microorganisms have reported the abundant presence of LPMOs, indicating their importance in biomass conversion [18,26]. However, the significance of LPMOs in starch utilization has remained unclear. Indeed, starch is regarded as a much less recalcitrant substrate compared to *e.g.* lignocellulose or chitin, indicating that LPMOs maybe not needed for efficient conversion. Nevertheless, two studies recently reported oxidative cleavage of starch by members of family AA13 LPMOs [15,16], indicating that LPMOs could be important for the degradation of this abundant polysaccharide. Indeed, LPMOs from both family AA9 and AA13 were identified in the starch secretomes, the latter family being the most abundant. As a matter of fact, one of the AA13 LPMOs, *An*LPMO13B (Q5B027) is amongst the top 4 most abundant proteins at all time points in the wheat and maize starch cultures (Fig. 4, Additional file 2: Tables S3−S4), indicating a prominent role in starch degradation. The other AA13 LPMO (*An*LPMO13A), which has a CBM20 appended, is also abundant in the secretomes of the cereal starches (always among the top 10%). The abundance of *An*LPMO13A may

be underestimated as compared to *An*LPMO13B, due to the preferential CBM20-mediated binding of the former onto granular starch (concluded form β-cyclodextrin capture, Fig. 5). *An*LPMO13A has already been demonstrated to cleave amylose chains and to boost hydrolysis of retrograded starch [16], whereas no activity data is available for *An*LPMO10B. However, the AA13 LPMO pair are highly similar (catalytic modules share 81% sequence identity), indicating that they also share substrate specificity. As already mentioned, *An*LPMO13A has a C-terminal CBM20. *An*LPMO13B, on the other hand, has a 75 amino acid C-terminal extension of unknown function. These differences in the two AA13 LPMOs may indicate that different parts of the starch structure are targeted, yielding complementary activities. Such a situation was observed by Forsberg *et al*., who showed that a conserved cellulose-targeting LPMO pair encoded by *Streptomyces coelicolor* gave synergy when combined in cellulose degradation reaction [41]. Interestingly no activity of the AA13 catalytic module lacking the CBM20 was detected [16]. This highlights the role of the CBM20 in targeting the enzyme to the insoluble starch structures.

Although less abundant than the family AA13 LPMOs, four AA9 LPMOs and one AA11 LPMOs were identified in the cultures (Table 1, Additional file 2: Tables S2). So far, AA9 LPMOs have only been shown to depolymerize cellulose and hemicellulose substrates. Thus, it is reasonable to speculate that these enzymes are secreted by the fungus to aid gaining access to the starch granules that *in vivo* are shielded by a protective lignocellulose layer, *e.g*. grain bran or to disrupt cell-walls in the starchy endosperm tissue. Interestingly, the five clade 3 AA9 LPMOs of the *A. nidulans* genome were not represented in the secretomes (Table 1). Clade 3 LPMOs are known to have mixed activity, yielding both C1 and C4 oxidized products, whereas clade 1 members oxidize the C1 carbon and clade 2 the C4 carbon of the substrate. The significance of this expression pattern is not known, but may be related to the type of substrate the fungus is exposed to.

The only AA11 LPMO observed in the secretomes is not abundant and the role of this enzyme is not obvious. The enzyme shows only moderate similarity to *Ao*AA11 (51%) [42], preventing prediction

of substrate specificity. Thus, it is difficult rationally explain why only one of the two is expressed when the fungus is grown on starch substrates. If the enzyme is active towards chitin (like *Ao*AA11), it is possible that the function is related to fungal cell wall remodeling

The role of non-LPMO redox active enzymes in starch degradation

It is well established that the subfamily 1 members of the glucose, methanol, choline (GMC) oxidoreductases (AA3s), *i.e*. the cellobiose dehydrogenases (CDHs), can provide electrons to LPMOs [25,43]. Maltodextrins are known as poor substrates for CDHs [44], thus it may be that other redox-active enzymes play this role during starch degradation. Indeed, multiple enzymes from the AA3 family and the AA7 family (glucooligosaccharide oxidases) are identified in the culture supernatant (Fig. 4). However, all AA3 and AA7 enzymes observed are classified as oxidases (not dehydrogenases as the CDHs), indicating that the main product of substrate oxidation is H_2O_2. On one hand, it is believed that such H_2O_2 generating enzymes are present to provide lignin degrading enzymes like laccases and peroxidases with hydrogen peroxide, which is a reaction component in the mechanisms of these lignin depolymerizing enzymes. On the other hand, no or little lignin is present in the substrates provided in the present experiments, possibly indicating a different role of these enzymes, but the presence of minor amounts of plant cell wall material that triggers the expression of lignin active enzymes cannot be precluded.

If AA3 and AA7 oxidases are active on products derived from starch depolymerization, thereby generating H_2O_2, the presence of Catalase B (CatB) as the dominant enzyme in most secretomes makes sense. Catalase catalyzing the conversion of H_2O_2 to O_2 and H_2O, protecting the surroundings from the potentially toxic effects of this compounds (*i.e*. hydroxyl radicals emerging from Fenton chemistry). Indeed, the expression of this CatB has previously been demonstrated to be induced by H_2O_2 or H_2O_2 -generating conditions [45], and the enzyme activity is suggested by the authors to protect the fungus from the toxic side-products encountered in aerobic growth. It should be noted that LPMOs, which represent some of the most abundant proteins in the cultures, are capable of

substantial H_2O_2 production [46]. A second highly abundant protein observed in the fungal cultures is thioredoxin reductase (Additional file 2: Tables S3–S5), a protein also related to detoxification of reactive oxygen species (ROS) [47]. Thus, it seems that the fungus is actively protecting itself from ROS during aerobic degradation of biomass. These findings could inspire future research on design of enzyme cocktails for biomass depolymerization, *i.e*. protecting enzymes from toxic byproducts formed by oxidative enzymes. As a matter of fact, one study has already shown the advantageous use of catalase in enzyme cocktails containing high levels of LPMOs [48].

AmyR regulation of amylolytic hydrolases and LPMOs

The AmyR transcription factor, responsive to isomaltose, has been shown to regulate the expression of the α-glucosidases agdA, agdB agdE agdF, the α-amylases amyA, amyB, AmyF, and the glucoamylase glaB in *A. nidulans* [49]. Most of these enzymes were indeed identified in the present study (Table 2). Analysis of the *A. nidulans* genome revealed the presence of the AmyR consensus sequence 5'-CGGN_8CGG-3' at around -300 basepairs (-318 to -305) in the promotor region of the *An*LPMO13A and around -600 basepairs (-612 to -609) for *An*LPMO13B. This is the first time a common regulatory link has been identified between GHs and LPMOs. Interestingly, hierarchical clustering analysis revealed that this enzyme has a similar secretion pattern as distinct amylolytic enzymes (Fig. 4). The abundance of both AA13 enzymes together with the modular glucoamylase GlcA and GH31 α-glucosidases increased in the last three days of the culture. This first data on the temporal distribution of starch degrading enzymes suggest a sophisticated regulatory mechanism, whereby the fungus deploys specific enzymes at different stages of the starch growth. It is tempting to speculate that the increased secretion of starch specific AA13 and other starch degrading enzymes correlates with the later stages of the culture, where the more resistant structures in the starch substrate accumulate. Further studies are needed to assess the rationale behind this specific secretion pattern.

CONCLUSIONS

This is the first study that uses high resolution secretomics analysis to probe the temporal distribution of protein profiles of the model saprophytic ascomycete *A. nidulans.* The data unveils a conspicuous abundance of AA13 LPMOs in the secretomes, which is suggestive of an instrumental role of these enzymes in starch degradation. Another novel finding is the identification of binding sites of the transcriptional regulator AmyR, which established a co-regulatory link between GHs and LPMOs featuring in starch degradation. Beyond a core amylolytic machinery, the secretomes were clearly correlated to the starch type used for growth and numerous CAZymes, proteases and oxidoreductases were secreted. A possible rationale for this is targeting non-starch component in the substrate matrix. The abundance and of LPMOs, AA3, AA7 and other oxidative enzymes is particularly interesting and merits further work to understand the role of these enzymes. This novel insight promotes our understanding of the degradation of different starches, including those of more resistant nature, and inspires formulation of better commercial enzyme cocktails for more efficient exploration of this important biomass resource.

MATERIALS AND METHODS

Carbohydrate substrates and assay chemicals

The wheat starch was from Sigma-Aldrich (S5127, unmodified), pea and high amylose (HA) maize starches were from KMC (Brande, Denmark). Starch substrates for the growth experiments were washed twice 70% ethanol and water and subsequently pelleted by centrifugation (14000 × *g,* 5 min) before resuspension in autoclaved media for growth experiments or as substrates for enzymatic assays. The starches were not autoclaved to minimize changes of the starch granule structure. *p*-Nitrophenyl-α-D-glucopyranoside (PNPG) and *p*-nitrophenol (PNP) were from Sigma-Aldrich, and insoluble Blue Starch was a custom preparation from Pharmacia (Uppsala, Sweden).

Fungal strain and culture conditions

A. nidulans strain FGSC A4 was obtained from the fungal genetics stock center (FGSC, Kansas City, MO). The fungus was pre-grown on malt extract agar plates containing 1% wheat, HA maize or pea starches for 5 days at 30°C until new mycelia were formed. Mycelial plugs were used to inoculate 1 L of minimal medium containing 1% (w/v) carbon source with a start pH of 6.5 in 3 L baffled shake flasks. The minimal medium contained, per liter, 6 g $NaNO_3$, 0.52 g KCl, 0.52 g $MgSO_4 \cdot 7H_2O$, 1.52 g KH_2PO_4, and 2 ml of Hutner's trace elements. Hutner's trace elements contained, per liter, 2.2 g $ZnSO_4 \cdot 7H_2O$, 1.1 g H_3BO_3, 0.5 g $MnCl_2 \cdot 4H_2O$, 0.5 g $FeSO_4 \cdot 7H_2O$, 0.16 g $CoCl_2 \cdot 6H_2O$, 0.16 g $CuSO_4 \cdot 5H_2O$, 0.11 g $(NH_4)Mo_7O_{24} \cdot 4H_2O$, and 5 g EDTA. The flasks were incubated on a rotary shaker (150 rpm) at 30°C for 5 days. All experiments were performed in biological triplicates. The medium was supplemented with antibiotics (100 µg/ml ampicillin and 34 µg/ml chloramphenicol) to inhibit possible bacterial contamination from the starch substrates.

α-Amylase activity

α-Amylase activity of the fungal cultures was assayed toward insoluble Blue Starch (iBS). The activity was measured using iBS (6.25 mg/ml) suspended in 10 mM MES, pH 6.5. The reaction mixture (900 µl) was incubated for 15 min at 37°C after addition of the culture filtrate (100 µl). The reaction was stopped by addition of 0.5 M NaOH (200 µl). After centrifugation (4000 × *g*, 3 min) 200 µl supernatant (in triplicates) was transferred to a 96 well microtiter plate and A_{620} values were used to determine enzyme activity. One activity unit (U) was defined as the amount of enzyme that leads to an increase in A_{620} of 1 absorbance unit in the reaction mixture under these experimental conditions.

α-Glucosidase activity

The α-glucosidase activity was determined using *p*-nitrophenyl α-D-glucopyranoside (PNPG) as the substrate. The activity in culture filtrates was assayed towards 2 mM PNPG in 10 mM MES, 0.005% Triton X-100, pH 6.5 and 10 µl culture supernatant in 50 µl reactions, at 37°C for 60 min. The

reaction was stopped by adding 1 M Na_2CO_3 (200 µl). The amount of released *p*-nitrophenol (PNP) was measured spectrophotometrically at A_{410} using PNP (0−2 mM) as a standard. One unit of activity (U) was defined as the amount of enzyme that released 1 µmol/min of PNP at the given conditions.

Preparation of secretome samples and mass spectrometry analysis

Culture filtrates containing the fungal secretomes were separated from mycelia and residual insoluble starch by centrifugation (15000 × *g*, 15 min, 4°C) and filtration using 0.45 µm hydrophilic PVDF membranes (Millipore) and stored at -20°C until further use. Samples (10 ml) from three biological replicate cultures were collected from day 3 to 5 under sterile conditions. Protein concentrations of the culture filtrates were determined using a Bradford assay (Bio-Rad Laboratories, Hercules, CA) with a BSA standard according to the manufacturer's instructions.

Soluble proteins of culture filtrates (2ml) were precipitated by direct addition of 500 µl cold 50% trichloroacetic acid (TCA), incubated at 4°C for 8 hours, centrifuged (15000 × *g*, 15 min, 4°C) to pellet precipitated protein and washed with 300 µl ice-cold 10 mM HCl and 90% acetone mix. After centrifugation (15000 × *g*, 15 min, 4°C), pellets were dried and dissolved in 100 µl 20 mM Tris-HCl at pH 8, reduced with 10 mM (final concentration) dithiotreitol (DTT) and incubated at room temperature for 30 min. Alkylation was performed by adding 15 mM (final concentration) iodoacetamide (IAA) and incubating at room temperature for 30 min in the dark. The proteins were digested with 20 µl 12.5 ng/µl sequencing-grade modified trypsin (Promega, Madison, WI) and incubated at 37°C for 16 hours. Trypsination was stopped with 0.5% (final concentration) trifluoroacetic acid (TFA). Samples were dried in SpeedVac and purified with ZipTip C18 pipette tips (Merck Millipore, Cork, Ireland) according to manufacturer's instructions. Purified samples were dried and dissolved in 2% acetonitrile (ACN) and 0.1% TFA mix.

The peptides were analyzed with two technical replicates using a nanoHPLC-MS/MS system consisting of a Dionex Ultimate 3000 RSLCnano (Thermo Scientific, Bremen, Germany) connected

to a Q-Exactive hybrid quadrupole-orbitrap mass spectrometer (Thermo Scientific) equipped with a nano-electrospray ion source. Samples were loaded onto a trap column (Acclaim PepMap100, C_{18}, 5 µm, 100 Å, 300 µm i.d. x 5 mm, Thermo Scientific) and backflushed onto a 50 cm analytical column (Acclaim PepMap RSLC C_{18}, 2 µm, 100 Å, 75 µm i.d., Thermo Scientific). At the start, the columns were in 96 % solution A (0.1 % (v/v) formic acid), 4% solution B (80% (v/v) ACN, 0.1% (v/v) formic acid). Peptides were eluted using a 90 min gradient developing from 4 to 13% (v/v) solution B in 2 minutes, 13 to 45% (v/v) in 70 minutes and finally to 55% B in 5 minutes before the wash phase at 90% B, all at a flow rate of 300 nl/min. In order to isolate and fragment the 10 most intense peptide precursor ions at any given time throughout the chromatographic elution, the Q-Exactive mass spectrometer was operated in data-dependent mode to switch automatically between orbitrap-MS and higher-energy collisional dissociation (HCD) orbitrap-MS/MS acquisition. The selected precursor ions were then excluded for repeated fragmentation for 20 seconds. The resolution was set to R=70,000 and R=35,000 for MS and MS/MS, respectively. For optimal acquisition of MS/MS spectra, automatic gain control (AGC) target values were set to 1,000,000 charges and a maximum injection time of 128 ms.

MS raw files were imported into MaxQuant [50,51] version 1.4.1.2 and proteins were identified and quantified using the MaxLFQ algorithm [52]. The samples were searched against a database containing the proteome of *A. nidulans* downloaded from UniProt (10557 sequences) [53] and supplemented with common contaminants such as keratins, trypsin and BSA. In addition, reversed sequences of all protein entries were concatenated to the database to estimate the false discovery rate (FDR). Protein N-terminal acetylation, oxidation of methionine, conversion of glutamine to pyro glutamic acid, and deamidation of asparagine and glutamine were used as variable modifications, while carbamidomethylation of cysteine residues was used as a fixed modification. Trypsin was used as digestion enzyme and two missed cleavages were allowed. The 'match between runs' feature of MaxQuant was enabled with default parameters, in order to transfer identifications

between samples based on accurate mass and retention time [52]. This was done to increase the number of identified peptides and was set so that transfer of peptides was only allowed between samples from the same substrate. All identifications were filtered in order to achieve a protein FDR of 1% and two ratio counts were required for a valid protein quantification.

Post-processing was done using Perseus version 1.5.0.31. Proteins categorized as only identified by site and matches to reversed sequences or contaminants were removed. Furthermore, the proteins were filtered so that a valid quantification existed for at least two of the three replicates on at least one substrate. Intensities were log-transformed and missing values imputed based on a downshifted normal distribution. Hierarchical clustering and heat map generation were done with Euclidean distance measure and average linkage. For visualization of trending proteins, ANOVA (permutation-based FDR, $p < 0.10$) was used to filter out proteins that showed no significant change over time. Protein quantitative values were z-scored within substrate and imputed values removed prior to clustering and heat map generation. Secretion prediction was a combination of SignalP [54], Phobius [55] and WoLF PSort [56] where at least two prediction algorithms had to agree. Secreted CAZymes were annotated using dbCAN (http://csbl.bmb.uga.edu/dbCAN/) [57].

β-Cyclodextrin affinity chromatography purification of secreted proteome and identification of most abundant proteins

The supernatant of the culture grown on wheat starch for 5 days was supplemented with $(NH_4)_2SO_4$ to 0.5 M final concentration and agitated until the salt was completely dissolved. Thereafter, the sample was filtrated through 0.45 µm membrane filters (Merck Millipore) and applied to an XK 16/20 column packed with a 20 ml β-cyclodextrin (β-CD) Sepharose affinity resin and pre-equilibrated with 4 column volumes of 10 mM sodium acetate buffer, pH 5.5, with 500 mM NaCl at 1 ml/min [58]. After sample loading at 0.75 ml/min, the column was washed with 8 column volumes of the above buffer at 1 ml/min. Bound proteins were eluted with 4 column volumes of 20 mM

sodium acetate buffer, 7 mM β-CD, pH 5.5. The purification was carried out using an ÄKTA Avant chromatograph interfaced by UNICORN 5.0 control software (GE Healthcare).

The eluted sample was analyzed by SDS-PAGE using Novex NuPAGE® 4-12% Bis-Tris Gels (Invitrogen, Carlsbad, CA) in the XCell *Sure*Lock Mini-Cell system (Invitrogen), according to manufacturer's instructions and visualized by staining with InstantBlue solution (Expedeon, Cambridgeshire, UK). Spots were excised from the visually most prominent protein bands from the stained gel, washed in 300 µl 40% ethanol at 50°C for 15 min and in 100 µl 100% ACN at room temperature for 10 min. The proteins in the gel were reduced with 50 µl 10 mM DTT in 100 mM NH_4HCO_3 at 56°C for 45 min. Cysteine alkylation was performed by adding 100 µl 55 mM IAA in 100 mM NH_4HCO_3 and incubating at room temperature for 30 min in the dark. The gels were washed with 100 µl 50% ACN, then 100 µl 100% ACN and dried. The proteins in the gel were digested with 10 µl 12.5 ng/µl sequencing-grade modified trypsin (Promega) dissolved in 10 mM NH_4HCO_3 and incubated at 4°C for 45 min, followed by addition of 10 µl 10 mM NH_4HCO_3 and incubation at 37°C for 16 hours. 1 or 2 µl samples were spotted directly onto a MTP AnchorChip target plate (Bruker Daltonics, Bremen, Germany), allowed to dry, and overlaid with 1 µl 0.5 µg/µl α-cyano-4-hydroxycinnamic acid (CHCA) matrix in 90% ACN, 0.1% TFA. The MS analyses were performed using an Ultraflex II MALDI-TOF/TOF MS (Bruker Daltonics). The obtained mass spectra were processed with FlexAnalysis and BioTools software both provided by the instrument manufacturer. Combination of MS and MS/MS data were used as input for databases searching for the spectra from MALDI-TOF/TOF using an in-house-licensed Mascot search engine (Matrix Science, London, UK). Proteins were identified using NCBInr database. The following parameters were set for searching: allowed global modification, carbamidomethyl cysteine; variable modification, oxidation of methionine; missed cleavages, 1; peptide tolerance, 80 ppm; MS/MS tolerance, ±0.5 Da. The protein identification was considered valid if it matched more than 2 peptides and the significance threshold for protein identifications was $p < 0.05$.

Identification of the consensus binding for the AmyR transcription factor

The AmyR consensus binding site sequence 5'-CGGN$_8$CGG-3' was identified by searching 1000 bp upstream of the start codon of the *A. nidulans* genes encoding *An*LPMO13A and *An*LPMO13B. The positions of the 5' and 3' ends of the AmyR binding site with putative translational start site as +1 were calculated.

Abbreviations

AA: auxiliary activity enzyme; CAZyme: Carbohydrate active enzyme; CBM: Carbohydrate binding module; CDH: Cellobiose dehydrogenase; FGSC: Fungal genetic stock center; GH: Glycoside hydrolase; HA: High amylose; LC-MS/MS: Liquid chromatography-tandem mass spectroscopy; LPMO: Lytic polysaccharide monooxygenase; PL: Pectin lyase, β-CD: β-Cyclodextrin.

Acknowledgements

This work was supported by a grant (NNF12OC0000769) from the Novo Nordisk foundation in the area "Biotechnology-based Synthesis and Production". GVK was funded by grant 214138 provided by the Norwegian Research Council.

Author's contribution

LN and MAH designed the experiments. LN performed the experiments. MØA performed secretome analysis. LN analyzed secretomics data. LN, MAH and GVK wrote the manuscript. All authors participated in finalizing the manuscript and have read and approved the final manuscript.

REFERENCES

1. Zeeman SC, Kossmann J, Smith AM. Starch: its metabolism, evolution, and biotechnological modification in plants. *Annu Rev Plant Biol*. 2010;61:209-234.
2. Pérez S, Bertoft E. The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review. *Starch/Staerke*. 2010;62(8):389-420.
3. Buléon A, Colonna P, Planchot V, Ball S. Starch granules: structure and biosynthesis. *Int J Biol Macromol*. 1998;23(2):85-112.
4. Tester RF, Karkalas J, Qi X. Starch - composition, fine structure and architecture. *J Cereal Sci*. 2004;39(2):151-165.
5. Sun H, Zhao P, Ge X, Xia Y, Hao Z, Liu J, Peng M. Recent advances in microbial raw starch degrading enzymes. *Appl Biochem Biotechnol*. 2010;160(4):988-1003.
6. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørlie M, Eijsink VGH. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*. 2010;330(6001):219-222.
7. Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JCN, Johansen KS, Krogh KBRM, Jørgensen CI, Tovborg M, Anthonsen A, et al. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci*. 2011;108(37):15079-15084.
8. Vu VV, Beeson WT, Phillips CM, Cate JHD, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc*. 2014;136(2):562-565.
9. Hemsworth GR, Davies GJ, Walton PH. Recent insights into copper-containing lytic polysaccharide mono-oxygenases. *Curr Opin Struct Biol*. 2013;23(5):660-668.
10. Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunæs AC, Stenstrøm Y, Mackenzie A, Sørlie M, Horn SJ, Eijsink VGH. Cleavage of cellulose by a CBM33 protein. *Protein Sci*. 2011;20(9):1479-1483.
11. Johansen KS. Discovery and industrial applications of lytic polysaccharide mono-oxygenases. *Biochem Soc Trans*. 2016;44(1):143-149.
12. Isaksen T, Westereng B, Aachmann FL, Agger JW, Kracher D, Kittl R, Ludwig R, Haltrich D, Eijsink VGH, Horn SJ. A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem*. 2014;289(5):2632-2642.
13. Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WGT, Ludwig R, Horn SJ, Eijsink VGH,

Westereng B. Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci*. 2014;111(17):6287-6292.

14. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SWA, Koetsier MJ, van Berkel WJH, Gruppen H, Kabel MA. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels*. 2015;8(1):101.
15. Vu VV, Beeson WT, Span EA, Farquhar ER, Marletta MA. A family of starch-active polysaccharide monooxygenases. *Proc Natl Acad Sci*. 2014;111(38):13822-13827.
16. Lo Leggio L, Simmons TJ, Poulsen JCN, Frandsen KEH, Hemsworth GR, Stringer MA, von Freiesleben P, Tovborg M, Johansen KS, De Maria L, et al. Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nat Commun*. 2015;6:5961.
17. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels*. 2013;6(1):41.
18. Navarro D, Rosso MN, Haon M, Olivé C, Bonnin E, Lesage-Meessen L, Chevret D, Coutinho PM, Henrissat B, Berrin JG. Fast solubilization of recalcitrant cellulosic biomass by the basidiomycete fungus *Laetisaria arvalisinvolves* successive secretion of oxidative and hydrolytic enzymes. *Biotechnol Biofuels*. 2014;7(1):143.
19. Saykhedkar S, Ray A, Ayoubi-Canaan P, Hartson SD, Prade R, Mort AJ. A time course analysis of the extracellular proteome of *Aspergillus nidulans* growing on sorghum stover. *Biotechnol Biofuels*. 2012;5(1):52.
20. Wortman JR, Gilsenan JM, Joardar V, Deegan J, Clutterbuck J, Andersen MR, Archer D, Bencina M, Braus G, Coutinho P, et al. The 2008 update of the *Aspergillus nidulans* genome annotation: a community effort. *Fungal Genet Biol*. 2009;46(1):S2-S13.
21. Book AJ, Yennamalli RM, Takasuka TE, Currie CR, Phillips GN, Fox BG. Evolution of substrate specificity in bacterial AA10 lytic polysaccharide monooxygenases. *Biotechnol Biofuels*. 2014;7(1):109.
22. Christiansen C, Abou Hachem M, Janecek S, Viksø-Nielsen A, Blennow A, Svensson B. The carbohydrate-binding module family 20 - diversity, structure, and function. *FEBS J*. 2009;276(18):5006-5029.
23. Ho DP, Ngo HH, Guo W. A mini review on renewable sources for biofuel. *Bioresour Technol*. 2014;169:742-749.

24. Yue D, You F, Snyder SW. Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. *Comput Chem Eng*. 2014;66:36-56.
25. Langston JA, Shaghasi T, Abbate E, Xu F, Vlasenko E, Sweeney MD. Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl Environ Microbiol*. 2011;77(19):7007-7015.
26. Hori C, Ishida T, Igarashi K, Samejima M, Suzuki H, Master E, Ferreira P, Ruiz-Duenas FJ, Held B, Canessa P, et al. Analysis of the *Phlebiopsis gigantea* genome, transcriptome and secretome provides insight into its pioneer colonization strategies of wood. *PLoS Genet*. 2014;10(12):e1004759.
27. Ring SG, Gee JM, Whittam M, Orford P, Johnson IT. Resistant starch: Its chemical form in foodstuffs and effect on digestibility *in vitro*. *Food Chem*. 1988;28(2):97-109.
28. Smith AM. The biosynthesis of starch granules. *Biomacromolecules*. 2001;2(2):335-341.
29. Ratnayake WS, Hoover R, Warkentin T. Pea starch: Composition, structure and properties — A review. *Starch/Staerke*. 2002;54(6):217-234.
30. Planchot V, Colonna P, Gallant DJJ, Bouchet B. Extensive degradation of native starch granules by α-amylase from *Aspergillus fumigatus*. *J Cereal Sci*. 1995;21(2):163-171.
31. Salgado P, Freire JPB, Ferreira RB, Teixeira A, Bento O, Abreu MC, Toullec R, Lalles JP. Immunodetection of legume proteins resistant to small intestinal digestion in weaned piglets. *J Sci Food Agric*. 2003;83(15):1571-1580.
32. Palmer R, Cornuault V, Marcus SE, Knox JP, Shewry PR, Tosi P. Comparative *in situ* analyses of cell wall matrix polysaccharide dynamics in developing rice and wheat grain. *Planta*. 2014;241(3):669-685.
33. Knudsen KEB. Fiber and nonstarch polysaccharide content and variation in common crops used in broiler diets. *Poult Sci*. 2014;93(9):2380-2393.
34. Andersen MR, Vongsangnak W, Panagiotou G, Salazar MP, Lehmann L, Nielsen J. A trispecies *Aspergillus* microarray: comparative transcriptomics of three *Aspergillus* species. *Pnas*. 2008;105(11):4387-4392.
35. Le DM, Fojan P, Azem E, Pettersson D, Pedersen NR. Visualization of the anticaging effect of Ronozyme WX xylanase on wheat substrates. *Cereal Chem*. 2013;90(5):439-444.
36. Choct M. Feed non-starch polysaccharides: Chemical structures and nutritional significance. *Feed Milling Int*. 1997;191:13-26.
37. Coutinho PM, Andersen MR, Kolenova K, VanKuyk PA, Benoit I, Gruben BS, Trejo-Aguilar B, Visser H, Solingen P Van, Pakula T. Post-genomic insights into the plant polysaccharide

degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. *Fungal Genet Biol*. 2009;46(1):S161-S169.

38. Bauer S, Vasu P, Persson S, Mort AJ, Somerville CR. Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. *Proc Natl Acad Sci*. 2006;103(30):11417-11422.
39. Matsubara T, Ammar Y Ben, Anindyawati T, Yamamoto S, Ito K, Iizuka M, Minamiura N. Degradation of raw starch granules by α-amylase purified from culture of *Aspergillus awamori* KT-11. *J Biochem Mol Biol*. 2004;37(4):422-428.
40. van der Kaaij RM, Janecek S, van der Maarel MJEC, Dijkhuizen L. Phylogenetic and biochemical characterization of a novel cluster of intracellular fungal α-amylase enzymes. *Microbiology*. 2007;153(12):4003-4015.
41. Forsberg Z, Mackenzie AK, Sorlie M, Rohr ÅK, Helland R, Arvai AS, Vaaje-Kolstad G, Eijsink VGH. Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc Natl Acad Sci*. 2014;111(23):8446-8451.
42. Hemsworth GR, Henrissat B, Davies GJ, Walton PH. Discovery and characterization of a new family of lytic polysaccharide monooxygenases. *Nat Chem Biol*. 2014;10(2):122-126.
43. Tan T, Kracher D, Gandini R, Sygmund C, Kittl R, Haltrich D, Hällberg BM, Ludwig R, Divne C. Structural basis for cellobiose dehydrogenase action during oxidative cellulose degradation. *Nat Commun*. 2015;6(May):7542.
44. Henriksson G, Johansson G, Pettersson G. A critical review of cellobiose dehydrogenases. *J Biotechnol*. 2000;78(2):93-113.
45. Calera J a., Sánchez-Weatherby J, López-Medrano R, Leal F. Distinctive properties of the catalase B of *Aspergillus nidulans*. *FEBS Lett*. 2000;475(2):117-120.
46. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels*. 2012;5(1):79.
47. Thon M, Al-Abdallah Q, Hortschansky P, Brakhage AA. The thioredoxin system of the filamentous fungus *Aspergillus nidulans*: Impact on development and oxidative stress response. *J Biol Chem*. 2007;282(37):27259-27269.
48. Scott BR, Huang HZ, Frickman J, Halvorsen R, Johansen KS. Catalase improves saccharification of lignocellulose by reducing lytic polysaccharide monooxygenase-associated enzyme inactivation. *Biotechnol Lett*. 2016;38:425-434.

49. Nakamura T, Maeda Y, Tanoue N, Makita T, Kato M, Kobayashi T. Expression profile of amylolytic genes in *Aspergillus nidulans*. *Biosci Biotechnol Biochem*. 2006;70(10):2363-2370.
50. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008;26(12):1367-1372.
51. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen J V., Mann M. Andromeda: A peptide search engine integrated into the MaxQuant environment. *J Proteome Res*. 2011;10(4):1794-1805.
52. Cox J, Hein MY, Luber CA, Paron I. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics*. 2014;13(9):2513-2526.
53. Consortium TU. UniProt: a hub for protein information. *Nucleic Acids Res*. 2014;43(D):204-212.
54. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 2011;8(10):785-786.
55. Käll L, Krogh A, Sonnhammer ELL. Advantages of combined transmembrane topology and signal peptide prediction-the Phobius web server. *Nucleic Acids Res*. 2007;35(W):429-432.
56. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. WoLF PSORT: Protein localization predictor. *Nucleic Acids Res*. 2007;35(W):585-587.
57. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. DbCAN: A web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res*. 2012;40(W):445-451.
58. Vester-Christensen MB, Abou Hachem M, Naested H, Svensson B. Secretory expression of functional barley limit dextrinase by *Pichia pastoris* using high cell-density fermentation. *Protein Expr Purif*. 2010;69(1):112-119.
59. Frandsen KEH, Simmons TJ, Dupree P, Poulsen JCN, Hemsworth GR, Ciano L, Johnston EM, Tovborg M, Johansen KS, von Freiesleben P, et al. The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. *Nat Chem Biol*. 2016:1-6.

FIGURE LEGENDS

Fig. 1 Estimation of amylolytic activities. Measurement of secreted α-amylase (A) and α-glucosidase (B) activities from *Aspergillus nidulans* grown on wheat (black), high-amylose maize (grey) or pea starch (white) for 5 days (see Materials and Methods). The data is presented as the mean ± standard deviation.

Fig. 2 Overview of the secreted proteins in *Aspergillus nidulans* starch cultures. The figure depicts the distribution of the secreted proteins from culture supernatants of *A. nidulans* grown on wheat, high-amylose maize or pea starch on culture days 3–5. Identified carbohydrate active enzymes (CAZymes) are assigned in the following categories: auxiliary activity (AA), carbohydrate esterase (CE), glycoside hydrolase (GH), polysaccharide lyase (PL). Secreted proteins lacking functionally characterized homologues were assigned as uncharacterized and proteases are shown to highlight their abundance. The remaining proteins belong to a variety of functional categories and are represented as "other" for clarity.

Fig. 3 Heat map comparison of expression patterns of CAZymes detected after 3–5 days growth of *Aspergillus nidulans* on minimal media supplemented with: high-amylose (HA) maize, pea, or wheat starch. The colors in the heat map indicate the label-free quantification (LFQ) intensity reported by MaxQuant ranging from $2x10^6$ (light green) to $4x10^{11}$ (light red). Missing values were imputed from a normal distribution located at the quantification limit. Glycoside hydrolases (GHs) are in red, auxiliary activities (AAs) in blue, polysaccharide lyases (PLs) in purple, carbohydrate esterases (CEs) in green, and other proteins in grey. Secretion prediction is a combination of SignalP, Phobius and WolfPSort where at least 2 algorithms had to agree.

Fig. 4 Heat map identical to Fig. 3, but filtered for amylolytic and AA enzymes detected after 3–5 days growth of *Aspergillus nidulans* on minimal media supplemented with: high-amylose (HA) maize, pea, or wheat starch. Highlighted in green are AA13s, in purple – AA9s and AA11s, in blue – AA3s and AA7s, and in red – proteins associated with classical starch degradation.

Fig. 5 Purification of starch binding enzymes from the secretome. β-Cyclodextrin-Sepharose affinity chromatography purification of the *Aspergillus nidulans* wheat starch culture supernatant harvested after five days (lane II). Three dominant bands were identified using mass spectrometry (MALDI-TOF MS, see Materials and Methods). The bands were identified as glucoamylase (1), α-amylase (2 and 2*-CBM20 domain) and LPMO (3). Molecular weight in kDa (lane I).

Tables

Table 1 Predicted secreted lytic polysaccharide monooxygenases encoded in the *Aspergillus nidulans* genome

Protein name	CBM	Protein family	Uniprot	Found[a]	Closest structural homologue (PDB ID; % sequence identify)	Clade[b]	Putative activity
*An*LPMO9A		AA9	Q5BAP2	Y	*Nc*PMO-2 (4EIR; 35%)	LPMO2	C4
*An*LPMO9B	CBM1	AA9	Q5BCX8	Y	*Ls*(AA9)A (2AFC; 52%)	N.A.	C4[c]
*An*LPMO9C		AA9	Q5AZ52	Y	*Pc*LPMO10E (4B5Q; 49%)	LPMO1	C1
*An*LPMO9D		AA9	Q5B8T4	Y	*Nc*LPMO9C (4D7U; 29%)	LPMO1	C1
*An*LPMO9E	CBM1	AA9	Q5AQA6	N	*Ta*LPMO10A (2YET; 61%)	LPMO3	C1/C4
*An*LPMO9F		AA9	Q5B6H0	N	*Ta*LPMO10A (2YET; 55%)	LPMO3	C1/C4
*An*LPMO9G		AA9	Q5BEI9	N	*Ta*LPMO10A (2YET; 67%)	LPMO3	C1/C4
*An*LPMO9H		AA9	Q5B7G9	N	*Ta*LPMO10A (2YET; 57%)	LPMO3	C1/C4
*An*LPMO9I		AA9	Q5AUY9	N	*Ta*LPMO10A (2YET; 64%)	LPMO3	C1/C4
*An*LPMO11A		AA11	Q5AU55	Y	*Ao*(AA11) (4MAH; 51%)	N.A.	C1[d]
*An*LPMO11B		AA11	Q5BFS8	N	*Ao*(AA11) (4MAH; 40%)	N.A.	C1[d]
*An*LPMO13A	CBM20	AA13	Q5B1W7	Y	*Ao*LPMO13 (4OPB; 70%)	N.A.	C1[e]
*An*LPMO13B		AA13	Q5B027	Y	*Ao*LPMO13 (4OPB; 72%)	N.A.	C1[e]

AA: Auxiliary activity, LPMO: Lytic polysaccharide monooxygenase. SignalP was used to predict secretion signals.

[a] Putative LPMOs identified in *A. nidulans* secretome are indicated (Y), and not identified (N).

[b]Based classification suggested by [21].

[c]Based on activity data obtained for the closest structural homologue by [59].

[d]Based on activity data obtained for the closest structural homologue by [42].

[e]Based on activity data obtained for the closest structural homologue [15,16].

Table 2 Predicted secreted enzymes assigned in CAZy families implicated in starch degradation by the *Aspergillus nidulans* genome

Protein (name)	CBM	Protein family	Uniprot	Identified in secretome[a]
α-amylase (AmyA)		GH13.1	G5EB45	Y
α-amylase (AmyB)	CBM20	GH13.1	G5EAT0	Y
α-amylase (AmyD)		GH13.1	Q5B822	N
α-amylase (AmyE)		GH13.1	Q5AZF6	N
α-amylase (AmyF)		GH13.1	Q5B7U2	Y
α-glucoamylase (GlaA)	CBM20	GH15	C8VLL3	Y
α-glucoamylase (GlyB)	CBM20	GH15	Q5AWC8	N
α-glucosidase (AgdA)		GH31	G5EB03	Y
α-glucosidase (AgdB)		GH31	G5EB11	Y
α-glucosidase (AgdC)		GH31	Q5AWI5	N
α-glucosidase (AgdE)		GH31	Q5BET9	Y
α-glucosidase		GH31	Q5AU13	N

GH: Glycoside hydrolase, SignalP was used to predict secretion signals.

[a] Putative starch-degrading proteins identified in *A. nidulans* secretome are indicated (Y), and not identified (N).

Table 3 Identification of β-cyclodextrin-Sepharose affinity chromatography purified secreted proteins in *Aspergillus nidulans* culture grown on wheat starch for 5 days

Protein/ module	Protein family	Uniprot	MW[a] (kDa)	*p*I[a]	Score[b]	Sequence coverage[b], %
LPMO/ CBM20	AA13	Q5B1W7	42	4.7	138	16
α-Amylase/ CBM20	GH13	G5EAT0	69	4.8	85	6
α-Glucoamylase/ CBM20	GH15	Q5AWC8	71	5.3	211	6

[a]Hypothetical molecular weight and *p*I of the proteins.

[b]The score and sequence coverage (%) are based on the genome sequence information. MASCOT scores >26 indicate identity or significance threshold ($p<0.05$).

FIGURES

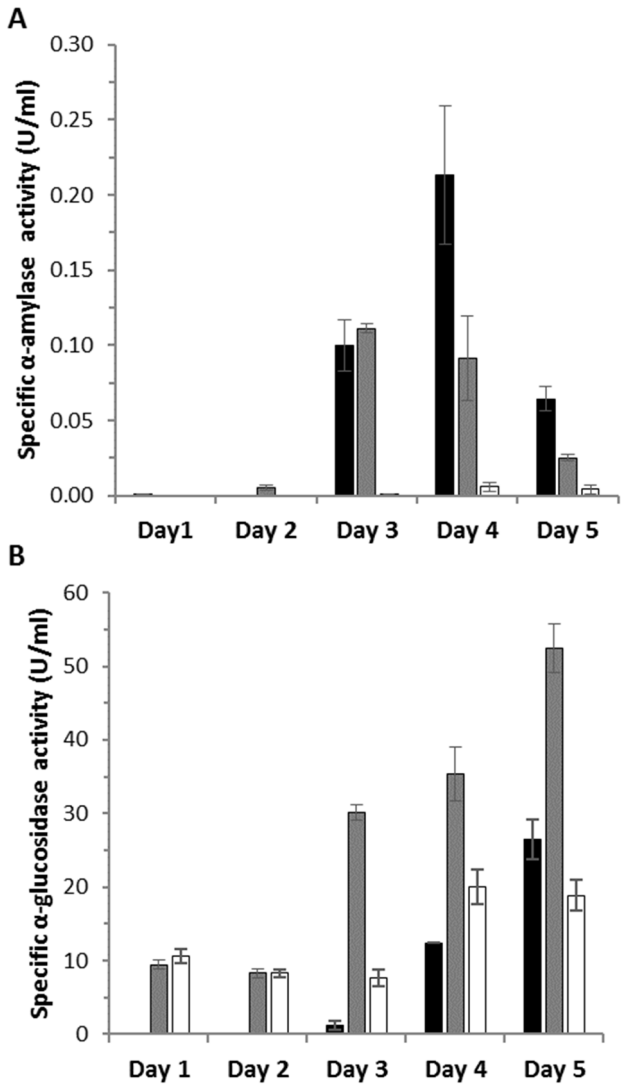


Fig. 1

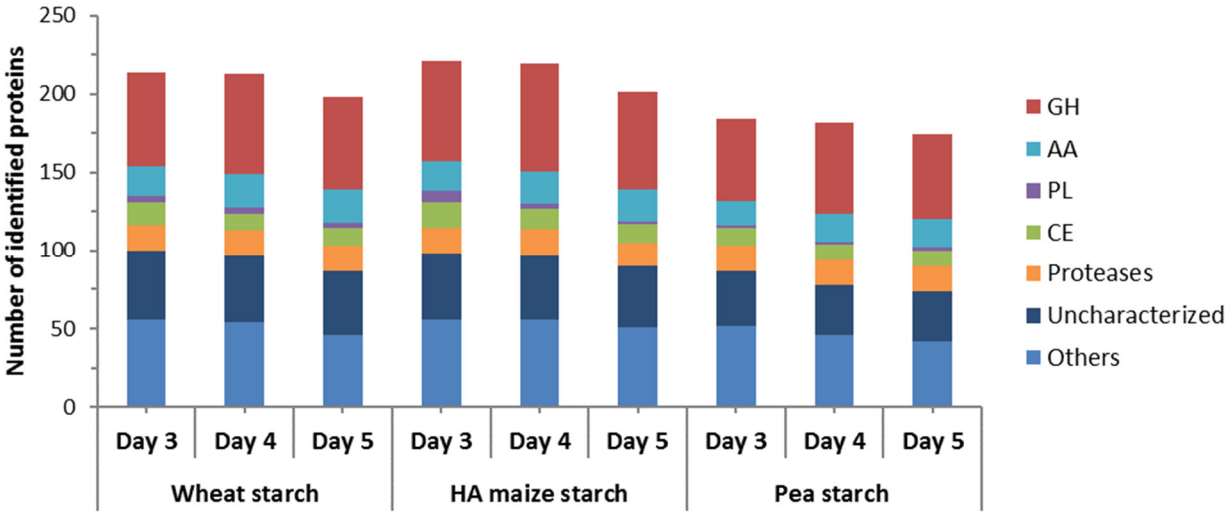

Number of identified proteins
250
200
150
100
50
0
Day 3
Day 4
Day 5
Day 3
Day 4
Day 5
Day 3
Day 4
Day 5
Wheat starch
HA maize starch
Pea starch
GH
AA
PL
CE
Proteases
Uncharacterized
Others

Fig. 2

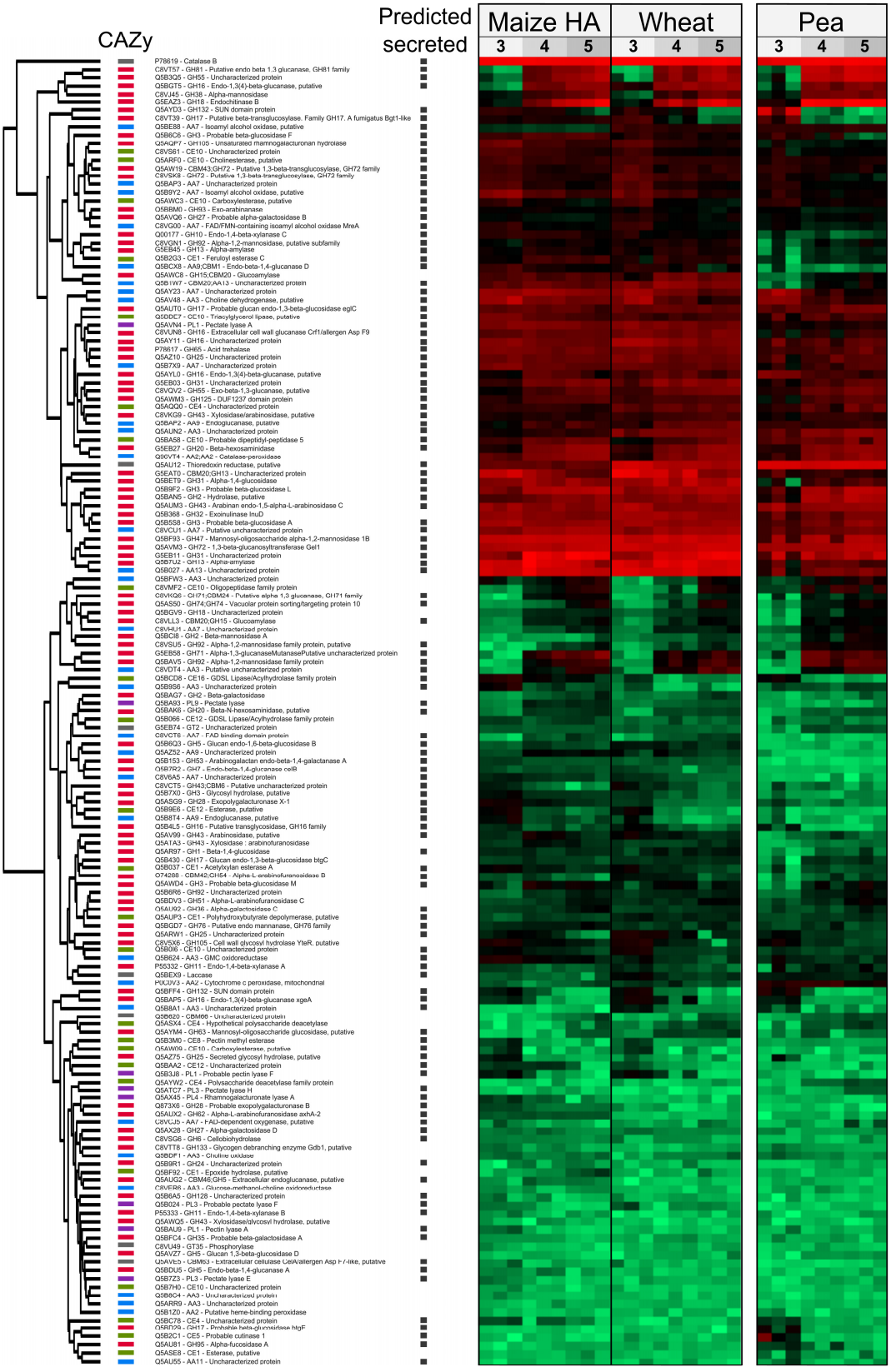

CAZy
Predicted secreted
Maize HA
Wheat
Pea
3
4
5
P78619 - Catalase B
C8VT57 - GH81 - Putative endo beta 1,3 glucanase, GH81 family
Q5B3Q5 - GH55 - Uncharacterized protein
Q5BGT5 - GH16 - Endo-1,3(4)-beta-glucanase, putative
C8VJ45 - GH38 - Alpha-mannosidase
G5EAZ3 - GH18 - Endochitinase B
Q5AYD3 - GH132 - SUN domain protein
C8VT39 - GH17 - Putative beta-transglucosylase. Family GH17. A fumigatus Bgt1-like
Q5BEB8 - AA7 - Isoamyl alcohol oxidase, putative
Q5B6C6 - GH3 - Probable beta-glucosidase F
Q5AQP7 - GH105 - Unsaturated rhamnogalacturonan hydrolase
C8VS61 - CE10 - Uncharacterized protein
Q5ARF0 - CE10 - Cholinesterase, putative
Q5AW19 - CBM43;GH72 - Putative 1,3-beta-transglucosylase, GH72 family
C8V5K8 - GH72 - Putative 1,3-beta-transglucosylase, GH72 family
Q5BAP3 - AA7 - Uncharacterized protein
Q5B9Y2 - AA7 - Isoamyl alcohol oxidase, putative
Q5AWC3 - CE10 - Carboxylesterase, putative
Q5BBM0 - GH93 - Exo-arabinanase
Q5AVQ6 - GH27 - Probable alpha-galactosidase B
C8VG00 - AA7 - FAD/FMN-containing isoamyl alcohol oxidase MreA
Q00177 - GH10 - Endo-1,4-beta-xylanase C
C8VGN1 - GH92 - Alpha-1,2-mannosidase, putative subfamily
G5EB45 - GH13 - Alpha-amylase
Q5B2G3 - CE1 - Feruloyl esterase C
Q5BCX8 - AA9;CBM1 - Endo-beta-1,4-glucanase D
Q5AWC8 - GH15;CBM20 - Glucoamylase
Q5B1W7 - CBM20;AA13 - Uncharacterized protein
Q5AY23 - AA7 - Uncharacterized protein
Q5AV48 - AA3 - Choline dehydrogenase, putative
Q5AUT0 - GH17 - Probable glucan endo-1,3-beta-glucosidase eglC
Q5DDC7 - CE10 - Triacylglycerol lipase, putative
Q5AVN4 - PL1 - Pectate lyase A
C8VUN8 - GH16 - Extracellular cell wall glucanase Crf1/allergen Asp F9
Q5AY11 - GH16 - Uncharacterized protein
P78617 - GH65 - Acid trehalase
Q5AZ10 - GH25 - Uncharacterized protein
Q5B7X9 - AA7 - Uncharacterized protein
Q5AYL0 - GH16 - Endo-1,3(4)-beta-glucanase, putative
G5EB03 - GH31 - Uncharacterized protein
C8VQV2 - GH55 - Exo-beta-1,3-glucanase, putative
Q5AWM3 - GH125 - DUF1237 domain protein
Q5AQG0 - CE4 - Uncharacterized protein
C8VKG9 - GH43 - Xylosidase/arabinosidase, putative
Q5BAP2 - AA9 - Endoglucanase, putative
Q5AUN2 - AA3 - Uncharacterized protein
Q5BAS8 - CE10 - Probable dipeptidyl-peptidase 5
G5EB27 - GH20 - Beta-hexosaminidase
Q9UVT4 - AA2;AA2 - Catalase-peroxidase
Q5AU12 - Thioredoxin reductase, putative
G5EAT0 - CBM20;GH13 - Uncharacterized protein
Q5BET9 - GH31 - Alpha-1,4-glucosidase
Q5B9F2 - GH3 - Probable beta-glucosidase L
Q5BAN5 - GH2 - Hydrolase, putative
Q5AUM3 - GH43 - Arabinan endo-1,5-alpha-L-arabinosidase C
Q5B368 - GH32 - Exoinulinase InuD
Q5BSS8 - GH3 - Probable beta-glucosidase A
C8VCU1 - AA7 - Putative uncharacterized protein
Q5BF93 - GH47 - Mannosyl-oligosaccharide alpha-1,2-mannosidase 1B
Q5AVM3 - GH72 - 1,3-beta-glucanosyltransferase Gel1
G5EB11 - GH31 - Uncharacterized protein
Q5B7U2 - GH13 - Alpha-amylase
Q5B027 - AA13 - Uncharacterized protein
Q5BFW3 - AA3 - Uncharacterized protein
C8VMF2 - CE10 - Oligopeptidase family protein
C8VKQ6 - GH71;CBM24 - Putative alpha 1,3 glucanase, GH71 family
Q5AS50 - GH74;GH74 - Vacuolar protein sorting/targeting protein 10
Q5BGV9 - GH18 - Uncharacterized protein
C8VLL3 - CBM20;GH15 - Glucoamylase
C8VHU1 - AA7 - Uncharacterized protein
Q5BCI8 - GH2 - Beta-mannosidase A
C8VSU5 - GH92 - Alpha-1,2-mannosidase family protein, putative
G5EB58 - GH71 - Alpha-1,3-glucanaseMutanasePutative uncharacterized protein
Q5BAV5 - GH92 - Alpha-1,2-mannosidase family protein
C8VDT4 - AA3 - Putative uncharacterized protein
Q5BCD8 - CE16 - GDSL Lipase/Acylhydrolase family protein
Q5BBS6 - AA3 - Uncharacterized protein
Q5BAG7 - GH2 - Beta-galactosidase
Q5BA93 - PL9 - Pectate lyase
Q5BAK6 - GH20 - Beta-N-hexosaminidase, putative
Q5B066 - CE12 - GDSL Lipase/Acylhydrolase family protein
G5EB74 - GT2 - Uncharacterized protein
C8VCT6 - AA7 - FAD binding domain protein
Q5B6Q3 - GH5 - Glucan endo-1,6-beta-glucosidase B
Q5AZS2 - AA9 - Uncharacterized protein
Q5B153 - GH53 - Arabinogalactan endo-beta-1,4-galactanase A
Q5B7R2 - GH7 - Endo-beta-1,4-glucanase celB
C8V6A5 - AA7 - Uncharacterized protein
C8VCT5 - GH43;CBM6 - Putative uncharacterized protein
Q5B7X0 - GH3 - Glycosyl hydrolase, putative
Q5ASG9 - GH28 - Exopolygalacturonase X-1
Q5B9E6 - CE12 - Esterase, putative
Q5B874 - AA9 - Endoglucanase, putative
Q5B4L5 - GH16 - Putative transglycosidase, GH16 family
Q5AV99 - GH43 - Arabinosidase, putative
Q5ATA3 - GH43 - Xylosidase : arabinofuranosidase
Q5ARS7 - GH1 - Beta-1,4-glucosidase
Q5B430 - GH17 - Glucan endo-1,3-beta-glucosidase btgC
Q5B037 - CE1 - Acetylxylan esterase A
O74288 - CBM42;CH54 - Alpha-L-arabinofuranosidase B
Q5AWD4 - GH3 - Probable beta-glucosidase M
Q5B6R6 - GH92 - Uncharacterized protein
Q5BDV3 - GH51 - Alpha-L-arabinofuranosidase C
Q5AU92 - GH36 - Alpha-galactosidase C
Q5AUP3 - CE1 - Polyhydroxybutyrate depolymerase, putative
Q5BGD7 - GH76 - Putative endo mannanase, GH76 family
Q5ARW1 - GH25 - Uncharacterized protein
C8V5X6 - GH105 - Cell wall glycosyl hydrolase YteR. putative
Q5B0I6 - CE10 - Uncharacterized protein
Q5B624 - AA3 - GMC oxidoreductase
P56332 - GH11 - Endo-1,4-beta-xylanase A
Q5BEX9 - Laccase
P0C0V3 - AA2 - Cytochrome c peroxidase, mitochondrial
Q5BFF4 - GH132 - SUN domain protein
Q5BAP5 - GH16 - Endo-1,3(4)-beta-glucanase xgeA
Q5B8A1 - AA3 - Uncharacterized protein
Q5B620 - CBM66 - Uncharacterized protein
Q5ASX4 - CE4 - Hypothetical polysaccharide deacetylase
Q5AYM4 - GH63 - Mannosyl-oligosaccharide glucosidase, putative
Q5B3M0 - CE8 - Pectin methyl esterase
Q5AW09 - CE10 - Carboxylesterase, putative
Q5AZ75 - GH25 - Secreted glycosyl hydrolase, putative
Q5BAA2 - CE12 - Uncharacterized protein
Q5B3J8 - PL1 - Probable pectin lyase F
Q5AYW2 - CE4 - Polysaccharide deacetylase family protein
Q5ATC7 - PL3 - Pectate lyase H
Q5AX45 - PL4 - Rhamnogalacturonate lyase A
Q873X6 - GH28 - Probable exopolygalacturonase B
Q5AUX2 - GH62 - Alpha-L-arabinofuranosidase axhA-2
C8VCJ5 - AA7 - FAD-dependent oxygenase, putative
Q5AX28 - GH27 - Alpha-galactosidase D
C8VSG6 - GH6 - Cellobiohydrolase
C8VTT8 - GH133 - Glycogen debranching enzyme Gdb1, putative
Q5BDF1 - AA3 - Choline oxidase
Q5B9R1 - GH24 - Uncharacterized protein
Q5BF92 - CE1 - Epoxide hydrolase, putative
Q5AUG2 - CBM46;GH5 - Extracellular endoglucanase, putative
C8VER6 - AA3 - Glucose-methanol-choline oxidoreductase
Q5B6A5 - GH128 - Uncharacterized protein
Q5B024 - PL3 - Probable pectate lyase F
P55333 - GH11 - Endo-1,4-beta-xylanase B
Q5AWQ5 - GH43 - Xylosidase/glycosyl hydrolase, putative
Q5BAU9 - PL1 - Pectin lyase A
Q5BFC4 - GH35 - Probable beta-galactosidase A
C8VU49 - GT35 - Phosphorylase
Q5AVZ7 - GH5 - Glucan 1,3-beta-glucosidase D
Q5AVE5 - CBM63 - Extracellular cellulase CelA/allergen Asp F7-like, putative
Q5BDU5 - GH5 - Endo-beta-1,4-glucanase A
Q5B7Z3 - PL3 - Pectate lyase E
Q5B7H0 - CE10 - Uncharacterized protein
Q5B8C4 - AA3 - Uncharacterized protein
Q5ARR9 - AA3 - Uncharacterized protein
Q5B1Z0 - AA2 - Putative heme-binding peroxidase
Q5BC78 - CE4 - Uncharacterized protein
Q5B7Z6 - GH17 - Probable beta-glucosidase btgF
Q5B2C1 - CE5 - Probable cutinase 1
Q5AU81 - GH95 - Alpha-fucosidase A
Q5ASE8 - CE1 - Esterase, putative
Q5AU55 - AA11 - Uncharacterized protein
GH
PL
Others
AA
CE

Fig. 3

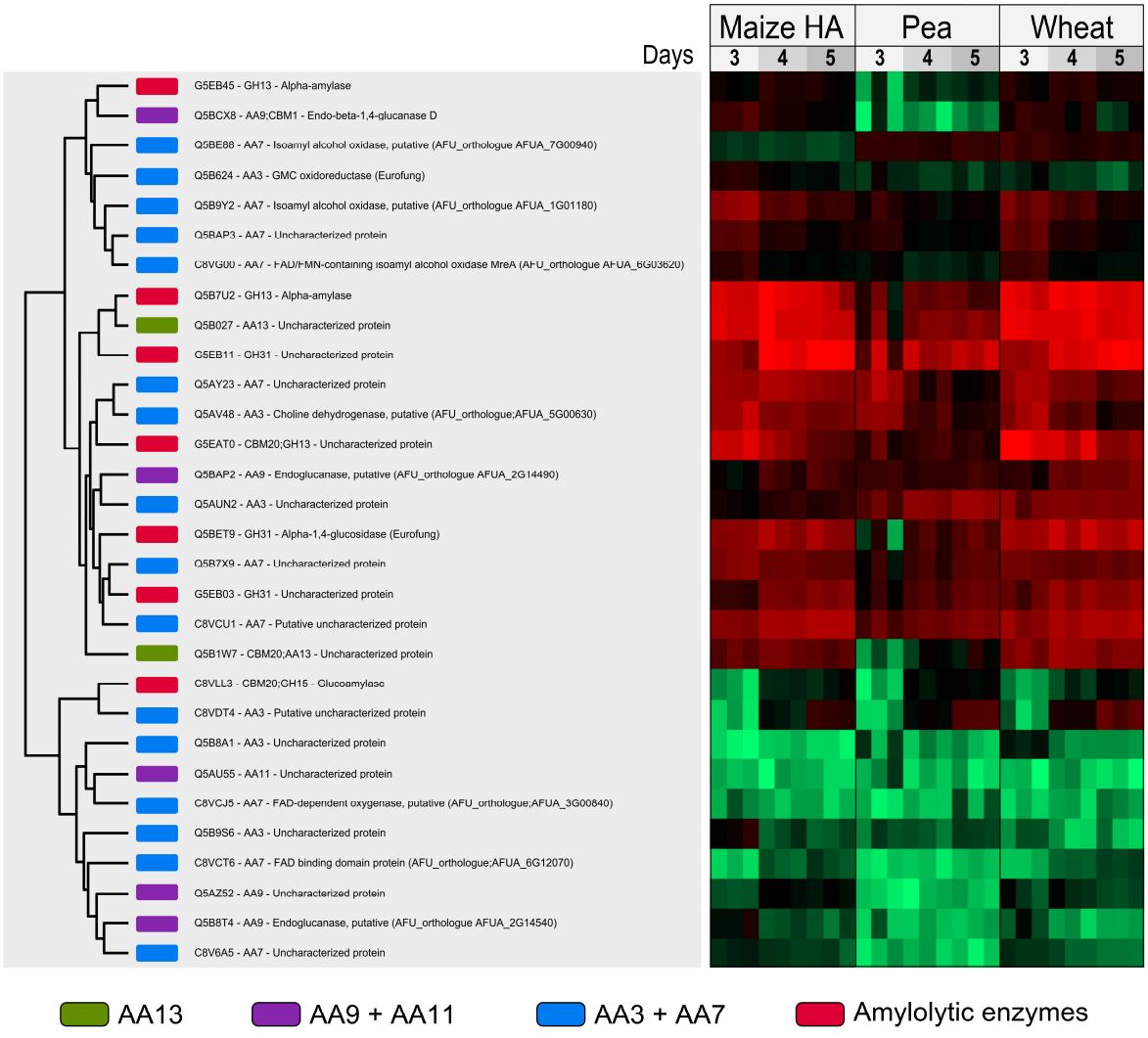

Days
Maize HA
Pea
Wheat
3 4 5 3 4 5 3 4 5
G5EB45 - GH13 - Alpha-amylase
Q5BCX8 - AA9;CBM1 - Endo-beta-1,4-glucanase D
Q5BE88 - AA7 - Isoamyl alcohol oxidase, putative (AFU_orthologue AFUA_7G00940)
Q5B624 - AA3 - GMC oxidoreductase (Eurofung)
Q5B9Y2 - AA7 - Isoamyl alcohol oxidase, putative (AFU_orthologue AFUA_1G01180)
Q5BAP3 - AA7 - Uncharacterized protein
C8VG00 - AA7 - FAD/FMN-containing isoamyl alcohol oxidase MreA (AFU_orthologue AFUA_6G03620)
Q5B7U2 - GH13 - Alpha-amylase
Q5B027 - AA13 - Uncharacterized protein
G5EB11 - GH31 - Uncharacterized protein
Q5AY23 - AA7 - Uncharacterized protein
Q5AV48 - AA3 - Choline dehydrogenase, putative (AFU_orthologue;AFUA_5G00630)
G5EAT0 - CBM20;GH13 - Uncharacterized protein
Q5BAP2 - AA9 - Endoglucanase, putative (AFU_orthologue AFUA_2G14490)
Q5AUN2 - AA3 - Uncharacterized protein
Q5BET9 - GH31 - Alpha-1,4-glucosidase (Eurofung)
Q5B7X9 - AA7 - Uncharacterized protein
G5EB03 - GH31 - Uncharacterized protein
C8VCU1 - AA7 - Putative uncharacterized protein
Q5B1W7 - CBM20;AA13 - Uncharacterized protein
C8VLL3 - CBM20;GH15 - Glucoamylase
C8VDT4 - AA3 - Putative uncharacterized protein
Q5B8A1 - AA3 - Uncharacterized protein
Q5AU55 - AA11 - Uncharacterized protein
C8VCJ5 - AA7 - FAD-dependent oxygenase, putative (AFU_orthologue;AFUA_3G00840)
Q5B9S6 - AA3 - Uncharacterized protein
C8VCT6 - AA7 - FAD binding domain protein (AFU_orthologue;AFUA_6G12070)
Q5AZ52 - AA9 - Uncharacterized protein
Q5B8T4 - AA9 - Endoglucanase, putative (AFU_orthologue AFUA_2G14540)
C8V6A5 - AA7 - Uncharacterized protein
AA13
AA9 + AA11
AA3 + AA7
Amylolytic enzymes

Fig. 4

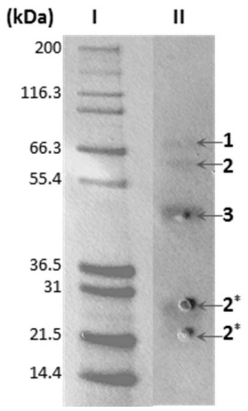

(kDa)
I
II
200
116.3
66.3
55.4
36.5
31
21.5
14.4
1
2
3
2*
2*

Fig. 5

Additional files

Additional file 1:

Table S1 Proteins identified in the secretome of *Aspergillus nidulans.* The table shows the MaxQuant output of identified proteins and the quantitative LFQ intensities associated with all replicates and conditions. Prediction of secretion was done as a combination of SignalP, Phobius and WoLF PSORT. CAZy annotation was done with dbCAN.

Additional file 2:

Table S2 Comparison of detected CAZymes in the secretome of *Aspergillus nidulans* during growth on wheat, high-amylose maize and pea starch at day 3, 4 and 5.

Table S3 Top 20 detected proteins in the secretome of *Aspergillus nidulans* during growth on wheat starch at day 3, 4 and 5.

Table S4 Top 20 detected proteins in the secretome of *Aspergillus nidulans* during growth on high amylose maize starch at day 3, 4 and 5.

Table S5 Top 20 detected proteins in the secretome of *Aspergillus nidulans* during growth on pea starch at day 3, 4 and 5.

Additional file 3:

Figure S1 Heat map comparison of expression patterns of 312 secreted proteins detected after 3–5 days growth of *A. nidulans* on minimal media supplemented with: high-amylose (HA) maize, pea, or wheat starch. The colors in the heat map indicate the label-free quantification (LFQ) intensity reported by MaxQuant ranging from $2x10^6$ (light green) to $4x10^{11}$ (light red). Missing values were imputed from a normal distribution located at the quantification limit. Highlighted in pink are AA13s, in light green – AA9s and AA11s, in dark green – AA3s and AA7s, in yellow – proteins associated with classical starch degradation, in brown – cell wall degrading enzymes, and in light brown – proteases. Secretion prediction is a combination of SignalP, Phobius and WolfPSort where at least 2 algorithms had to agree.

Figure S2 A trending heat map of proteins predicted to be secreted during growth of *Aspergillus nidulans* on high-amylose (HA) maize, pea or wheat starch. Two main clusters are indicated in the graph, proteins with decreasing levels over time (I) and proteins with increasing levels over time (II). The intensities are z-score normalized per row, separately for each substrate, to emphasize the trend of protein amount during growth independent of total level and substrate-level differences. Proteins showing no change over time (ANOVA, permutation based FDR > 0.1) were removed from the plot. Missing values (protein level less than detection limit of mass spectrometer) are colored white. Highlighted in pink are AA13s, in light green – AA9s and AA11s, in dark green – AA3s and AA7s, in yellow – proteins associated with classical starch degradation, in brown – cell wall degrading enzymes, and in light brown – proteases.

Additional File 2: Supplementary Tables S2 to S5

Supplementary Table S2: Comparison of detected CAZymes in the secretome of *Aspergillus nidulans* during growth on wheat, high-amylose maize and pea starch at day 3, 4 and 5

CAZy family		**Wheat starch**			**HA maize starch**			**Pea starch**		
		Day 3	*Day 4*	*Day 5*	*Day 3*	*Day 4*	*Day 5*	*Day 3*	*Day 4*	*Day 5*
GH1	PCW [(a)]	1	1	1	1	1	1	0	1	1
GH2	PCW	1	1	1	1	1	1	1	1	1
GH3	PCW	5	5	5	5	5	5	4	4	5
GH5	PCW	1	1	1	2	3	2	0	0	0
GH6	PCW	1	1	1	1	1	1	0	0	0
GH7	PCW	1	1	1	1	1	1	0	0	0
GH10	PCW	1	1	1	1	1	1	1	1	1
GH11	PCW	1	1	0	1	1	0	0	1	1
GH13	S	3	3	3	3	3	3	3	3	3
GH15	S	1	1	1	1	1	1	0	1	1
GH16		6	6	5	6	6	5	6	4	4
GH17		2	2	2	2	2	2	3	2	2
GH20		1	2	2	2	2	2	1	2	2
GH24		0	0	0	0	1	1	0	0	0
GH25		3	3	2	3	3	1	3	3	2
GH27	PCW	2	2	2	2	2	2	1	2	2
GH28	PCW	1	1	1	2	2	2	1	1	1
GH31	S	3	3	3	3	3	3	3	3	3
GH35	PCW	0	0	0	1	1	0	0	1	1
GH36	PCW	1	1	1	1	1	1	1	1	1
GH43	PCW	4	4	4	4	4	3	3	4	3
GH47		1	1	1	1	1	1	1	1	1
GH53	PCW	1	1	1	1	1	1	0	0	0
GH54	PCW	1	1	1	1	1	1	0	1	1
GH55		2	2	2	2	2	2	2	2	2
GH62	PCW	1	1	1	1	1	1	0	0	0
GH63		1	1	1	0	0	1	1	1	1
GH65		1	1	1	1	1	1	1	1	1
GH71		0	2	2	1	2	2	1	2	2
GH72		3	3	3	3	3	3	3	3	3
GH74	PCW	1	1	1	1	1	1	1	1	1
GH76		1	1	1	1	1	2	1	2	1
GH81		1	1	1	1	1	1	1	1	1
GH92		2	3	3	2	3	3	3	3	3

CAZy family		Wheat starch			HA maize starch			Pea starch		
		Day 3	*Day 4*	*Day 5*	*Day 3*	*Day 4*	*Day 5*	*Day 3*	*Day 4*	*Day 5*
GH93	PCW	1	1	1	1	1	1	1	1	1
GH95	PCW	0	0	0	0	0	0	1	1	0
GH105	PCW	1	1	1	1	1	1	1	1	1
GH125		1	1	1	1	1	1	1	1	1
GH132		2	2	0	2	2	1	2	2	0
AA3		5	5	5	4	5	5	6	6	6
AA7		8	10	10	9	10	9	7	8	8
AA9	PCW	4	4	4	4	4	4	1	1	1
AA11		0	0	0	0	0	0	1	0	0
AA13	S	2	2	2	2	2	2	1	2	2
PL1	PCW	2	2	1	3	1	1	1	1	1
PL3	PCW	1	1	1	2	0	0	0	1	1
PL4	PCW	0	0	0	1	1	0	0	0	0
PL9	PCW	1	1	1	1	1	1	0	0	0
CE1	PCW	3	3	3	3	3	3	2	3	3
CE4	PCW	1	1	1	2	1	1	2	1	1
CE5		0	0	0	0	0	0	1	0	0
CE8	PCW	1	1	1	1	0	0	0	0	0
CE10		7	6	6	7	6	6	6	6	6
CE12		2	1	1	2	2	1	0	0	0
CE16	PCW	1	0	1	1	1	1	1	0	0

AA: Auxiliary Activity, CE: Carbohydrate Esterase, GH: Glycoside Hydrolase, GT: Glycoside Transferase, PL: Polysaccharide Lyase. (a): The families potentially active on plant cell walls (PCW) and starch (S) are indicated.

Supplementary Table S3: Top 20 detected proteins in the secretome of *Aspergillus nidulans* during growth on wheat starch at day 3, 4 and 5

	Protein	Protein family	Uniprot
Day 3			
1	Catalase B		P78619
2	Uncharacterized protein	GH13/CBM20	G5EAT0
3	Serine protease similarity, trypsin family		Q5BAR4
4	Uncharacterized protein	AA13	Q5B027
5	Thioredoxin reductase, putative		Q5AU12
6	α-Amylase	GH13	Q5B7U2
7	Aminopeptidase Y, putative		Q5ATD5
8	β-1,3-Glucanosyltransferase Gel1	GH72	Q5AVM3
9	Alkaline protease 1		Q00208
10	Uncharacterized protein		C8V6E2
11	Mannosyl-oligosaccharide α-1,2-mannosidase 1B	GH47	Q5BF93
12	Putative uncharacterized protein		Q5AWZ9
13	Extracellular serine-rich protein, putative		Q5B926
14	Arabinan endo-α-1,5-L-arabinosidase C	GH43	Q5AUM3
15	Uncharacterized protein		Q5AYU5
16	Uncharacterized protein	GH16	Q5AY11
17	Neutral protease 2 homolog		Q5AUR8
18	Probable β-glucosidase A	GH3	Q5B5S8
19	Cell wall mannoprotein MnpA		C8VP91
20	Choline dehydrogenase, putative	AA3	Q5AV48
Day 4			
1	Catalase B		P78619
2	α-Amylase	GH13	Q5B7U2
3	Uncharacterized protein	AA13	Q5B027
4	Aminopeptidase Y, putative		Q5ATD5
5	Uncharacterized protein	GH31	G5EB11
6	Uncharacterized protein		C8V6E2
7	Uncharacterized protein	GH13/CBM20	G5EAT0
8	Mannosyl-oligosaccharide α-1,2-mannosidase 1B	GH47	Q5BF93
9	Serine protease similarity, trypsin family		Q5BAR4
10	Alkaline protease 1		Q00208
11	Neutral protease 2 homolog		Q5AUR8
12	β-1,3-Glucanosyltransferase Gel1	GH72	Q5AVM3
13	α-1,4-Glucosidase	GH31	Q5BET9
14	Extracellular serine-rich protein, putative		Q5B926
15	Thioredoxin reductase, putative		Q5AU12

	Protein	Protein family	Uniprot
16	Uncharacterized protein		Q5AYU5
17	Putative uncharacterized protein		Q5AWZ9
18	Putative uncharacterized protein	AA7	C8VCU1
19	Probable β-glucosidase L	GH3	Q5B9F2
20	Glutaminase A		C8VAK7
Day 5			
1	Catalase B		P78619
2	Uncharacterized protein	GH31	G5EB11
3	β-1,3-Glucanosyltransferase Gel1	GH72	Q5AVM3
4	Uncharacterized protein	AA13	Q5B027
5	α-Amylase	GH13	Q5B7U2
6	Aminopeptidase Y, putative		Q5ATD5
7	Mannosyl-oligosaccharide α-1,2-mannosidase 1B	GH47	Q5BF93
8	Alkaline protease 1		Q00208
9	Uncharacterized protein		C8V6E2
10	Uncharacterized protein		Q5AYU5
11	Extracellular serine-rich protein, putative		Q5B926
12	α-1,4-Glucosidase	GH31	Q5BET9
13	Putative uncharacterized protein		Q5AWZ9
14	Probable β-glucosidase L	GH3	Q5B9F2
15	Thioredoxin reductase, putative		Q5AU12
16	Neutral protease 2 homolog		Q5AUR8
17	Putative uncharacterized protein	AA7	C8VCU1
18	Uncharacterized protein	GH55	Q5B3Q5
19	Serine protease similarity, trypsin family		Q5BAR4
20	Putative endo β-1,3-glucanase	GH81	C8VT57

AA: Auxiliary Activity, GH: Glycoside Hydrolase.

Supplementary Table S4: Top 20 detected proteins in the secretome of *Aspergillus nidulans* during growth on high amylose maize starch at day 3, 4 and 5

	Protein	Protein family	Uniprot
Day 3			
1	Catalase B		P78619
2	Uncharacterized protein	AA13	Q5B027
3	Uncharacterized protein	GH13/CBM20	G5EAT0
4	Aminopeptidase Y, putative		Q5ATD5
5	α-Amylase	GH13	Q5B7U2
6	Mannosyl-oligosaccharide α-1,2-mannosidase 1B	GH47	Q5BF93
7	Alkaline protease 1		Q00208
8	Putative uncharacterized protein		Q5AWZ9
9	β-1,3-Glucanosyltransferase Gel1	GH72	Q5AVM3
10	Serine protease similarity, trypsin family		Q5BAR4
11	Choline dehydrogenase, putative	AA3	Q5AV48
12	Extracellular serine-rich protein, putative		Q5B926
13	Neutral protease 2 homolog		Q5AUR8
14	Uncharacterized protein		C8V6E2
15	Uncharacterized protein		Q5B9G2
16	Probable β-glucosidase A	GH3	Q5B5S8
17	Arabinan endo-α-1,5-L-arabinosidase C	GH43	Q5AUM3
18	Uncharacterized protein		Q5AYU5
19	Isoamyl alcohol oxidase, putative	AA7	Q5B9Y2
20	Uncharacterized protein	AA7	Q5AY23
Day 4			
1	Catalase B		P78619
2	α-Amylase	GH13	Q5B7U2
3	Uncharacterized protein	GH31	G5EB11
4	Uncharacterized protein	AA13	Q5B027
5	Neutral protease 2 homolog		Q5AUR8
6	Mannosyl-oligosaccharide α-1,2-mannosidase 1B	GH47	Q5BF93
7	Aminopeptidase Y, putative		Q5ATD5
8	Alkaline protease 1		Q00208
9	β-1,3-Glucanosyltransferase Gel1	GH72	Q5AVM3
10	Probable β-glucosidase L	GH3	Q5B9F2
11	Uncharacterized protein		C8V6E2
12	Glutaminase A		C8VAK7
13	Extracellular serine-rich protein, putative		Q5B926
14	Putative uncharacterized protein	AA7	C8VCU1
15	Uncharacterized protein	GH13/CBM20	G5EAT0

	Protein	Protein family	Uniprot
16	Uncharacterized protein		Q5AYU5
17	Uncharacterized protein	AA7	Q5AY23
18	Putative uncharacterized protein		Q5AWZ9
19	α-1,4-Glucosidase	GH31	Q5BET9
20	Serine protease similarity, trypsin family		Q5BAR4
Day 5			
1	Uncharacterized protein	GH31	G5EB11
2	Catalase B		P78619
3	Uncharacterized protein	AA13	Q5B027
4	Mannosyl-oligosaccharide α-1,2-mannosidase 1B	GH47	Q5BF93
5	β-1,3-glucanosyltransferase Gel1	GH72	Q5AVM3
6	α-Amylase	GH13	Q5B7U2
7	Alkaline protease 1		Q00208
8	Probable β-glucosidase L	GH3	Q5B9F2
9	Uncharacterized protein	GH55	Q5B3Q5
10	Putative uncharacterized protein		Q5AWZ9
11	Putative uncharacterized protein	AA7	C8VCU1
12	Uncharacterized protein		C8V6E2
13	Extracellular serine-rich protein, putative		Q5B926
14	Aminopeptidase Y, putative		Q5ATD5
15	Neutral protease 2 homolog		Q5AUR8
16	Putative endo β 1,3 glucanase	GH81	C8VT57
17	Glutaminase A		C8VAK7
18	β-Hexosaminidase	GH20	G5EB27
19	Probable β-glucosidase A	GH3	Q5B5S8
20	α-1,4-Glucosidase	GH31	Q5BET9

AA: Auxiliary Activity, GH: Glycoside Hydrolase.

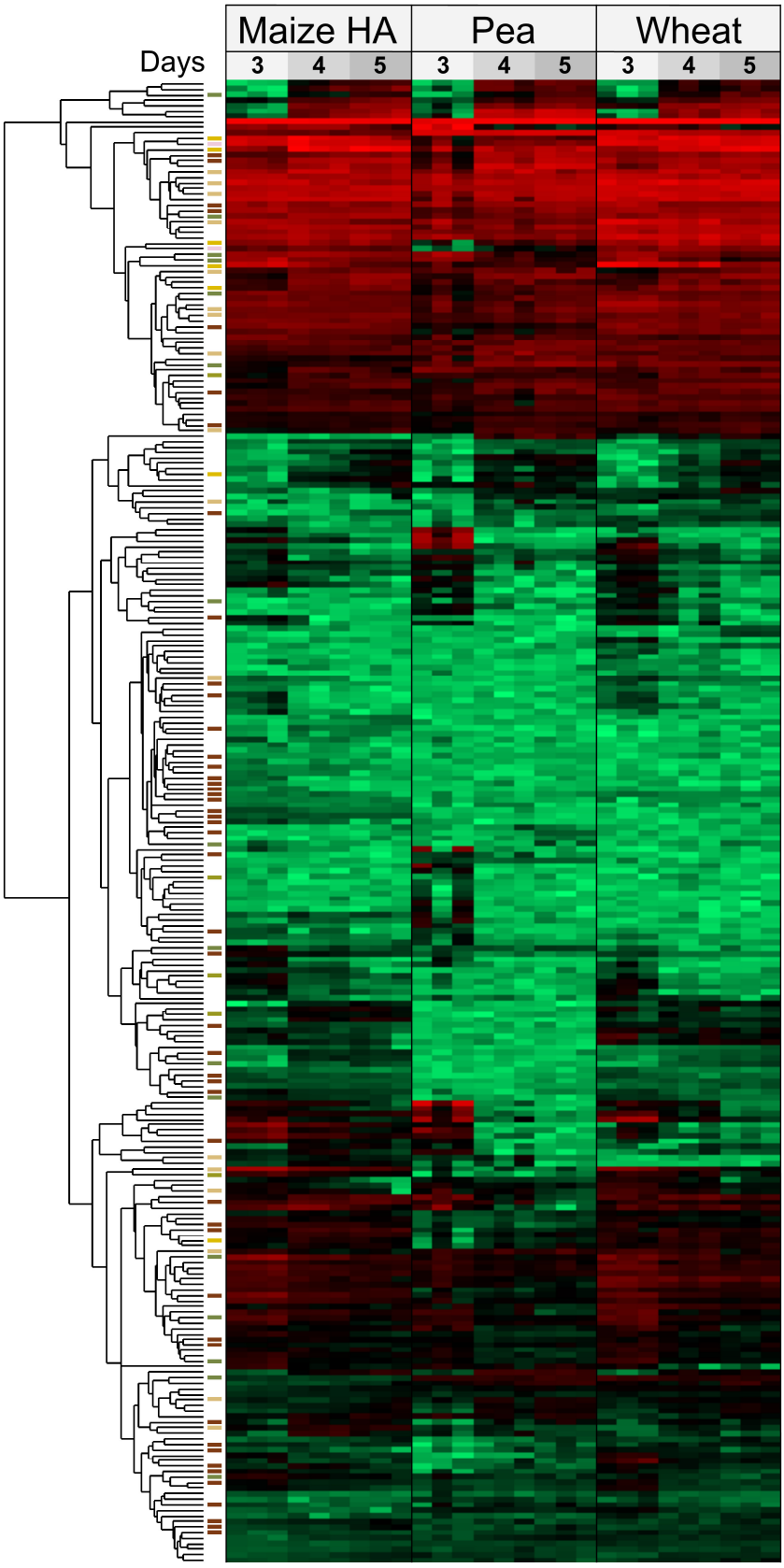
Supplementary Table S5: Top 20 detected proteins in the secretome of *Aspergillus nidulans* during growth on pea starch at day 3, 4 and 5

	Protein	Protein family	Uniprot
Day 3			
1	Uncharacterized protein		Q5B4F5
2	Catalase B		P78619
3	Thioredoxin reductase, putative		Q5AU12
4	SUN domain protein	GH132	Q5AYD3
5	Serine protease similarity, trypsin family		Q5BAR4
6	Aminopeptidase Y, putative		Q5ATD5
7	Putative uncharacterized protein		Q5AWZ9
8	Uncharacterized protein	AA7	Q5AY23
9	Uncharacterized protein		C8V6E2
10	Uncharacterized protein	GH31	G5EB11
11	β-1,3-Glucanosyltransferase Gel1	GH72	Q5AVM3
12	Arabinan endo-α-1,5-L-arabinosidase C	GH43	Q5AUM3
13	Allergen Asp F7		Q5B501
14	Uncharacterized protein		Q5AT80
15	Alkaline protease 1		Q00208
16	Choline dehydrogenase, putative	AA3	Q5AV48
17	Uncharacterized protein		Q5B277
18	Cell wall mannoprotein MnpA		C8VP91
19	Mannosyl-oligosaccharide α-1,2-mannosidase 1B	GH47	Q5BF93
20	Uncharacterized protein	AA7	Q5AY23
Day 4			
1	Catalase B		P78619
2	Thioredoxin reductase, putative		Q5AU12
3	Glutaminase A		C8VAK7
4	Aminopeptidase Y, putative		Q5ATD5
5	Neutral protease 2 homolog		Q5AUR8
6	Putative endo-β-1,3-glucanase	GH81	C8VT57
7	Uncharacterized protein	GH55	Q5B3Q5
8	Alkaline protease 1		Q00208
9	Uncharacterized protein	GH31	G5EB11
10	Putative uncharacterized protein		Q5AWZ9
11	Hydrolase, putative	GH2	Q5BAN5
12	β-1,3-Glucanosyltransferase Gel1	GH72	Q5AVM3
13	Mannosyl-oligosaccharide α-1,2-mannosidase 1B	GH47	Q5BF93
14	Probable β-glucosidase L	GH3	Q5B9F2
15	Arabinan endo-α-1,5-L-arabinosidase C	GH43	Q5AUM3
16	Hypothetical serine protease		C8VUL6

	Protein	Protein family	Uniprot
17	Uncharacterized protein	AA3	Q5AUN2
18	Serine protease similarity, trypsin family		Q5BAR4
19	Extracellular serine-rich protein, putative		Q5B926
20	β-Hexosaminidase	GH20	G5EB27
Day 5			
1	Catalase B		P78619
2	Thioredoxin reductase, putative		Q5AU12
3	Glutaminase A		C8VAK7
4	Uncharacterized protein	GH55	Q5B3Q5
5	Putative endo β-1,3-glucanase	GH81	C8VT57
6	Uncharacterized protein	GH31	G5EB11
7	Probable β-glucosidase L	GH3	Q5B9F2
8	Alkaline protease 1		Q00208
9	β-1,3-glucanosyltransferase Gel1	GH72	Q5AVM3
10	Aminopeptidase Y, putative		Q5ATD5
11	Mannosyl-oligosaccharide α-1,2-mannosidase 1B	GH47	Q5BF93
12	Hydrolase, putative	GH2	Q5BAN5
13	Putative uncharacterized protein		Q5AWZ9
14	Extracellular serine-rich protein, putative		Q5B926
15	Neutral protease 2 homolog		Q5AUR8
16	Arabinan endo-α-1,5-L-arabinosidase C	GH43	Q5AUM3
17	Uncharacterized protein	AA3	Q5AUN2
18	β-Hexosaminidase	GH20	G5EB27
19	Probable β-glucosidase A	GH3	Q5B5S8
20	Uncharacterized protein		C8V6E2

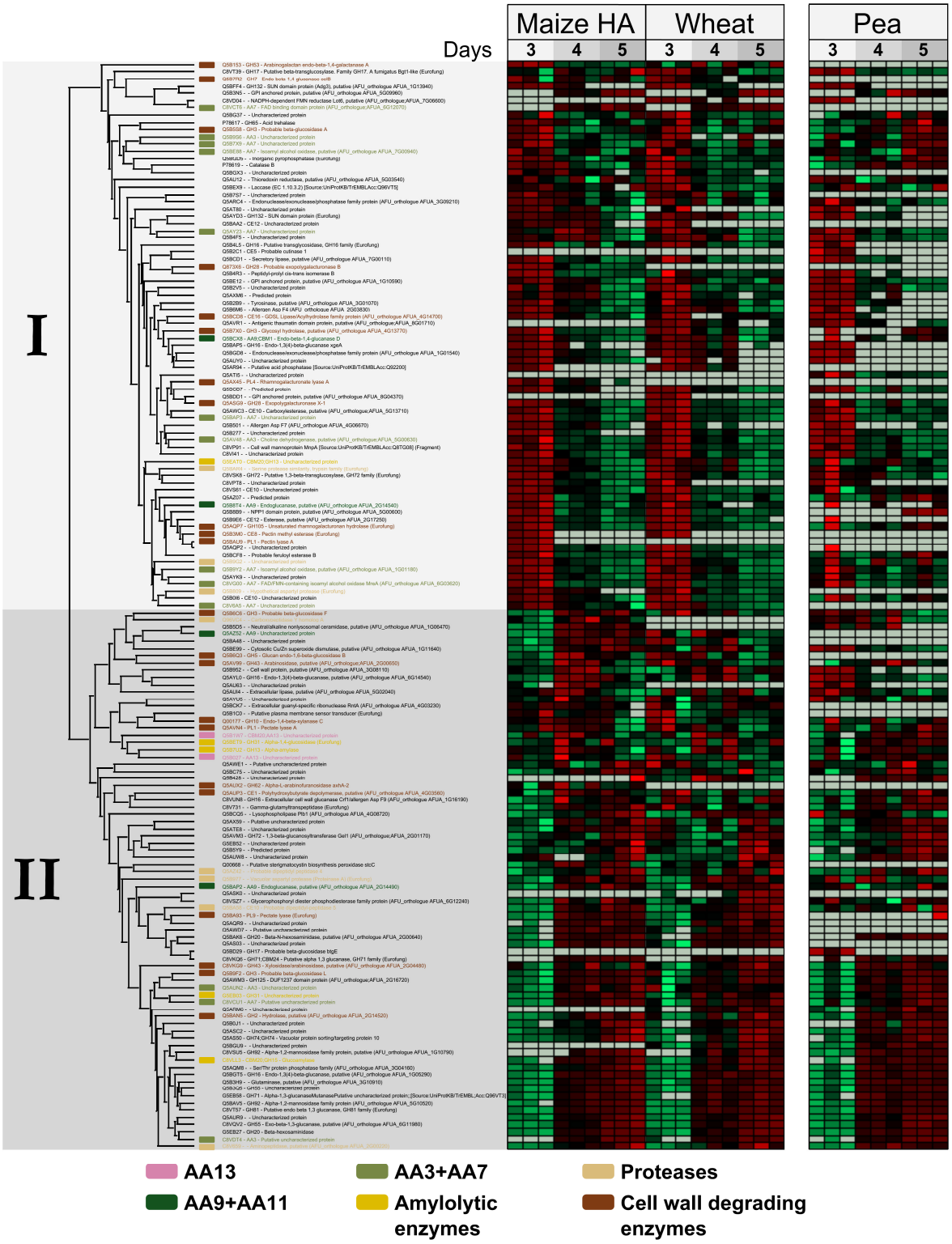
AA: Auxiliary Activity, GH: Glycoside Hydrolase.

Additional File 3: Supplementary Figures S1 and S2



AA13
AA9+AA11
AA3+AA7
Amylolytic enzymes
Proteases
Cell wall degrading enzymes

Add. file 3: Figure S1



Additional file 3: Figure S2

CHAPTER 3. PAPER II

A starch specific fungal lytic polysaccharide monooxygenase binds to starch with similar affinity as amylolytic hydrolases

A starch specific fungal lytic polysaccharide monooxygenase binds to starch with similar affinity as amylolytic hydrolases

Laura Nekiunaite[1], Trine Isaksen[2], Gustav Vaaje-Kolstad[2] and Maher Abou Hachem[1]

1 Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark

2 Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, PO Box 5003, 1432 Ås, Norway

Correspondence

M. Abou Hachem, Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, Elektrovej, Building 375, DK-2800 Kgs. Lyngby, Denmark, Fax: +45 4588 6307, Tel: +45 4525 2732

E-mail: maha@bio.dtu.dk

Abstract

Starch-binding modules of family 20 (CBM20) are present in 60% of the lytic polysaccharide monooxygenases (LPMOs) that oxidatively cleave glucosidic bonds in resistant starch. This frequent occurrence highlights CBM20-mediated substrate binding as an important feature in starch-active LPMOs, binding data for which is currently lacking. This study shows similar binding properties of two recombinant fungal LPMOs as compared to CBM20-containing amylolytic hydrolases. The conservation of functionally-important residues and taxonomy-based clustering of CBM20s joint to starch-active LPMOs and hydrolases suggest that these starch binding modules have been retained in the evolution of hydrolytic and oxidative starch degrading activities.

Keywords: AA13; carbohydrate binding module; CBM20; lytic polysaccharide monooxygenase; starch binding; β-cyclodextrin.

Abbreviations

AA, auxiliary activity; CAZy, carbohydrate-active enzymes; CBM, carbohydrate-binding module; DSC, differential scanning calorimetry; GH, glycoside hydrolase; ITC, isothermal titration calorimetry; LPMO, lytic polysaccharide monooxygenase; SBS, starch-binding site; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; β-CD, β-cyclodextrin.

Introduction

Starch is a major renewable energy storage polysaccharide in plants and an important resource not only as a food, but also as an industrial feedstock in biofuels, pharmaceuticals, detergents, and cosmetics [1–3]. Starch consists of two types of homo-glucose polymers: the mainly linear α-1,4-linked amylose and amylopectin, constituting 65−82% (w/w) of the starch granule and differing from amylose by having a larger molecular mass and roughly 5% α-1,6-branches of 12−15 glucosyl units long on average [2,4,5]. Starch is biosynthesized as insoluble granules, varying in size, morphology, crystal packing and crystallinity that ranges from 15 to 45% depending on botanical origin [6–8]. Radially alternating amorphous and semi-crystalline layers in the starch granule arise from the packing of double helices formed by adjacent branches in amylopectin, with the semi-crystalline regions contributing to resistance of starch to enzymatic degradation [9,10]. Many industrial applications require the disruption of starch granules through hydrothermal, harsh chemical or enzymatic treatments [11–13]. Despite development of relatively efficient α-amylases and other starch degrading enzymes, there is still a significant margin for improving starch hydrolysis yields and shortening processing time, which would significantly reduce energy and costs of the process [14,15].

Typically, glycoside hydrolases (GHs) that degrade complex polysaccharides possess carbohydrate binding modules (CBMs) that promote enzyme-substrate proximity and thereby enhance catalytic efficiency [16,17]. Moreover, CBMs can also modulate the specificity and activity of cognate enzymes against plant cell walls polysaccharides [18,19]. Similarly, starch-specific CBMs, which are assigned into CBM20 in the CAZy database [20,21], are reported to potentate the activity of fungal amylolytic enzymes, *e.g*. α-amylases and glucoamylases on granular starch [22,23]. Catalytic efficiency gains have been also conferred to amylolytic enzymes by fusion of CBM20s, attesting the importance of these ancillary modules in the deconstruction of starch [24,25].

LPMOs are copper-dependent enzymes that use molecular oxygen and an external electron donor to cleave glycosidic bonds in various polysaccharides, such as cellulose [26,27], hemicelluloses [28,29], and chitin [30]. These enzymes are assigned into auxiliary activity (AA) families 9, 10, 11, and 13 in the CAZy database. Similar to other complex polysaccharide active enzymes, CBMs occur frequently (~30%) together with LPMO catalytic modules [31]. Knowledge on CBMs occurring with LPMOs, however, is scarce. Recently, modular LPMOs comprising a AA13 catalytic module joint to C-terminal starch-binding module CBM20 have been shown to be active on starch [32,33]. No activity,

however, could be demonstrated for the truncated *Aspergillus oryzae* enzyme (*Ao*AA13) lacking the CBM20 module, which is present in most enzymes from this family. The purification of AA13 enzymes using amylose affinity columns demonstrates their affinity to starchy ligands, but currently there is no data on their binding properties.

In this study, we have analyzed the binding of two AA13 enzymes from *Aspergillus terreus* and the cereal pathogen *Magnaporthe oryzae* to starch and the model ligand β-cyclodextrin which is commonly used as a starch mimic substrate. Our data establish the ability of starch-active AA13 LPMOs to bind starch granules and β-cyclodextrin with comparable affinities to typical amylolytic hydrolases, which highlights the common and important function of CBMs in granular starch degradation in both types of enzymes.

Materials and Methods

Cloning, production and purification of recombinant enzymes

The genes (cDNA) encoding lytic polysaccharide monooxygenases of auxiliary activity 13 (AA13) family from *M. oryzae, Mo*LPMO13A (UniProt: Q2KEQ8), and from *A. terreus*, *At*LPMO13A (UniProt: Q0CGA6) were synthesized by GenScript (Piscataway, NJ). The synthetic *Mo*LPMO13A gene was inserted into the pPICZα A vector (Invitrogen, Carlsbad, CA) using BstBI and XbaI sites, and *At*LPMO13A using XhoI and XbaI sites. These plasmids were linearized with PmeI and transformed into electrocompetent *Pichia pastoris* X-33 cells (Invitrogen) by electroporation, and selected on YPDS plates supplemented with 100 µg/ml zeocin following manufacturer's instructions (EasySelect™ Pichia Expression Kit; Invitrogen). Transformants were screened for protein production in BMGY medium containing 1% (v/v) glycerol and best secreting transformants were used to produce *Mo*LPMO13A and *At*LPMO13A in a 5 L Biostat B bioreactor (Sartorius Stedim Biotech, Goettingen, Germany) as previously described [34]. The culture supernatants were recovered by centrifugation (14 000 × *g*, 45 min, 4 °C) and filtered using 0.45 µm membrane filters (Millipore, Bedford, MA). Both proteins were purified by affinity chromatography using β-cyclodextrin (β-CD) sepharose, followed by size exclusion chromatography. The cell-free supernatant was supplemented with $(NH_4)_2SO_4$ to 0.5 M, followed by centrifugation (15 000 × *g*, 15 min, 4 °C), and re-filtration before loading onto a 20 ml β-CD sepharose [35] column and purification as previously described [36]. The elution fractions containing the purified protein were pooled, concentrated and loaded onto a HiLoad 16/60 Superdex G-75 size exclusion column (GE Healthcare, Uppsala, Sweden) and eluted in 10 mM Na-Acetate, 150

mM NaCl, pH 5.5 at 1 ml/min. Chromatographic steps were performed using an ÄKTA Explorer chromatograph (GE Healthcare) at 4 °C. Protein purity was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS−PAGE), and the fractions containing pure protein were pooled and concentrated with Amicon Ultra centrifugal filters (MWCO 10 kDa, Millipore). Protein concentrations were determined by absorbance A_{280}, using the theoretical extinction coefficients calculated using ExPASy server (*Mo*LPMO13A: 71360 $M^{-1}cm^{-1}$, and *At*LPMO13A: 82360 $M^{-1}cm^{-1}$) [37].

Deglycosylation

N-glycosylation sites were predicted using the servers NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) and GlycoEP (http://www.imtech.res.in/raghava/glycoep/). *O*-glycosylation sites were predicted using the servers NetOGlyc (http://www.cbs.dtu.dk/services/NetOGlyc/) and GlycoEP.

*At*LPMO13A (10 µg) was incubated with 2 µl Endo H (New England Biolabs, Ipswitch, MA) in a 20 µl reaction volume at 37 °C for 4 hours. *Mo*LPMO13A (1 µg) was incubated with 1 µl Jack bean α-mannosidase (Sigma-Aldrich, St. Louis, MO) and/or 1 µl β-mannosidase (Megazyme, Bray, Ireland) in a 10 µl reaction volume for 2 hours at 25 °C and for another 2 hours at 37 °C. Mobility shifts were analyzed using SDS-PAGE.

Thermal stability of the recombinant LPMOs

Differential scanning calorimetry (DSC) was used to analyze the conformational stabilities of *Mo*LPMO13A and *At*LPMO13A using a Nano DSC instrument (TA Instruments, New Castle, DE). Protein samples (10 µM) were dialyzed against 3 x 1000 volumes of 10 mM Na Acetate buffer, pH 5.5 for 24 hours, degassed and loaded into sample cells, and scanned (20−100 °C, 1 °C/min) with the dialysis buffer in the reference cell. Baseline scans, collected with buffer in both reference and sample cells, were subtracted from sample scans, and the Universal Analysis software (TA Instruments) with a DSC add-on was used to model the reference cell and baseline corrected thermograms using a two state scaled model to determine T_m (unfolding temperature, defined as the temperature of maximum apparent heat capacity) and the calorimetric heat of unfolding ΔH_{cal}.

Insoluble starch binding assay

Binding of *Mo*LPMO13A to insoluble wheat starch (Sigma-Aldrich) was analyzed using wheat starch, which was washed three times with MilliQ water, followed by 10 mM Na-Acetate, pH 5.5. Enzyme aliquots (2.3 µM or 4.7 µM) were added to 0, 1, 5, 10, 20, 30,

or 45 mg/ml starch suspensions in the above buffer to a final volume of 300 µl. The suspensions were incubated at 4 °C for 1 hour, and then centrifuged (4 000 × *g*, 5 min, 4 °C) to pellet the starch. Free enzyme concentrations (A_{280}) in the supernatants from each suspension were measured and used to determine the fraction of bound protein. A one site binding model was fit to the binding isotherms, where *P* is protein, *S* is starch and B_{max} is the maximum binding capacity: $[P_{bound}]=B_{max}\cdot[S]/(K_d+[S])$, by non-linear regression (Langmuir isotherm) using Prism 6 software (GraphPad, La Jolla, CA). Controls without added enzyme, and with enzyme and 2 mM β-CD were performed.

Isothermal titration calorimetry (ITC)

ITC measurements were carried out using a MicroCal iTC200 calorimeter (MicroCal, Northampton, MA). *At*LPMO13A (200 µM), dialyzed against 10 mM MES, pH 6.5 overnight, was titrated with 2 mM β-CD dissolved in the same buffer at 25 °C with an initial injection of 0.4 µl, followed by 21 injections of 1.8 µl. Baseline measurements were made using an identical injection regime in the absence of protein. A one binding site model was fit to the integrated normalized data to determine the binding parameters using the MicroCal Origin software package (OriginLab, Northampton, MA).

Bioinformatics analysis

The CAZy database was used to retrieve AA13 sequences. The characterized *Aspergillus nidulans* LPMO, *An*AA13 (UniProt: Q5B1W7), was used as a query in a BLAST search against the non-redundant protein database [38] to retrieve additional AA13 orthologues having CBM20 domain that were not assigned in the CAZy database. Sequences with alignment score above 200 were retrieved. Conserved domain searches were performed using NCBI Conserved Domain Database [39], and sequences without putative CBM20 were omitted. Multiple sequence alignments of AA13 catalytic modules and CBM20 domains from all retrieved AA13 orthologues were performed using MUSCLE [40] and rendered using ESPript [41]. Sequences of 24 CBM20s from fungal GH13 α-amylases and GH15 glucoamylases were used for comparison with 75 sequences of CBM20s from LPMOs. A phylogenetic tree was rendered from the alignment using the ClustalW2 phylogeny with default settings [40] and visualized in Dendroscope 3.5 [42].

Results and Discussion

Heterologous expression of the recombinant LPMOs and enzyme stability

*Mo*LPMO13A and *At*LPMO13A were produced and purified at high yields (Fig. 1). Both enzymes migrated as smeary bands with higher apparent molecular masses than the theoretically calculated from the sequences (42734 Da and 37220 Da for *At*LPMO13A and *Mo*LPMO13A, respectively). The larger observed size is likely a result of *O*- and/or *N*-glycosylations of the proteins, which are predicted for both LPMOs. Particularly, *O*-glycosylation in the serine/threonine-rich linker, connecting the catalytic module to the C-terminal CBM20 (Fig. S1). Moreover, *At*LPMO13A has a predicted *N*-glycosylation site (N379) in the CBM20 (Fig. S1), at a position suggested to interact with the substrate at starch binding site 1 [23]. Glycosylation of this site is likely to reduce affinity due to steric hindering. Treatment with EndoH, which cleaves *N*-glycans did not result in visible change of migration pattern on the gels (data not shown) concluding that the discrepancy is the apparent size of *At*LPMO13A is mainly attributed to *O*-glycosylation. A combination of α- and β-mannosidase treatment of the LPMOs resulted in an apparent decrease in size on SDS-PAGE gels (data not shown), which confirms *O*-mannosylation of the enzymes as observed for other LPMOs expressed in *P. pastoris* [43–45]. The yield of mannosidases treatment was too low to allow a preparative deglycosylation of the enzyme.

The stability of the LPMOs was investigated using DSC analysis, which showed that both enzymes were highly thermostable attesting their structural integrity. The denaturing temperatures (T_m) were 70.0 and 70.9 °C, for *At*LPMO13A and *Mo*LPMO13A, respectively (Fig. S2). Both thermograms featured a single peak suggesting that the unfolding processes of the catalytic module and the CBM20 were overlapping. These T_m temperatures were slightly higher but comparable to other reported T_m values for fungal LPMOs (63.0−68.9 °C) [45]. This is also in agreement with the large calorimetric enthalpy of approximately 480 kJ/mol for both proteins.

Starch-specific LPMOs bind to starch granules and the starch mimic β-cyclodextrin

*Mo*LPMO13A and *At*LPMO13A were purified using β-cyclodextrin (β-CD) affinity chromatography suggesting that both enzymes possess affinity for starch. Indeed, the binding of *Mo*LPMO13A to wheat starch granules was demonstrated (Fig. 2). Control

experiments performed in the presence of 2 mM β-CD abolished binding, which confirms the binding specificity and precludes aggregation artefacts (data not shown). An equilibrium dissociation constant K_d=8.70±1.90 mg/ml and a maximum binding capacity of 35% were determined from the binding isotherms. This affinity is roughly one order of magnitude lower than that reported for the binding of the CBM20 from the cyclodextrin glycosyltransferase from *Bacillus circulans* using potato starch as ligand (K_d=0.79 mg/ml) [46]. This enzyme, however, possess significant affinity in its active site to β-cyclodextrin and helical structures in starch. By comparison, the affinity of the *Mo*LPMO13A to wheat starch was roughly one order higher than the enzyme comprising the homologous CBM20 from the *Aspergillus niger* glucoamylase fused to a β-galactosidase using corn starch as ligand (K_d=55.6 mg/ml) [25]. Use of different starches and experimental conditions, obviously affect the size of the binding constants. *M. oryzae* is a rice pathogen [47,48], while *A. terreus* for example is frequently associated with crops *e.g*. wheat and potatoes [49,50]. The binding of *Mo*LPMO13A to wheat starch, however, clearly demonstrates comparable binding to amylolytic CBM20-containing hydrolases.

β-CD is widely accepted as a model probe for starch binding and binding data for different starch-binding proteins are reported for this ligand [51–53]. The binding of the recombinant *At*LPMO13A to this model ligand was measured using ITC and the thermogram and binding isotherm are shown in Fig. 3. The data established moderate affinity binding to β-CD with an association equilibrium constant, K_a=(2.74±0.18)x10^4 M^{-1} equivalent to a dissociation constant, K_d=37 μM and a favorable free energy change, ΔG=-25.4 kJ/mol. The binding was driven largely by a favorable enthalpy (ΔH=-63.0±0.48 kJ/mol), which was off-set by an unfavorable entropy (-TΔS=37.6 kJ/mol). The binding affinity of *At*LPMO13A to β-CD is slightly lower, but in the same range as the K_d (≈10 μM) obtained for the isolated CBM20 from the *A. niger* glucoamylase [54] and the K_d (≈19 μM) for the full length enzyme [55]. Moreover, the thermodynamic signature of binding is also very similar, which suggests that the binding affinity of the CBM20 appended to the *A. terreus* LPMO is similar to counterparts that occur with starch degrading hydrolases. Notably, the obtained binding stoichiometry (*n*) was 0.54 as compared to value of two obtained for typical CBM20 domains that possess two binding sites [55]. It cannot be excluded that the binding stoichiometry is reduced due to the presence of a fraction of the enzymes with impaired binding due to steric hindrance caused by glycosylation (Fig. S1), but the ITC affinity and thermodynamic parameters clearly demonstrate a typical binding pattern of modular GHs that target starch aided by a CBM20 module. It is not possible from this data to rule out the presence of additional binding sites on the catalytic module, which

may be impaired due to glycosylation. Interestingly, activity was only possible to demonstrate on the full length LPMO from *A. nidulans*, which suggests that the starch binding mediated by the CBM20 contributes importantly to the potency of starch active LPMOs [32]. Solvent accessible aromatic residues are almost invariably present in the active site of GHs, but are conspicuously lacking in the active sites of LPMOs, which is indicative of a more dynamic lower affinity substrate binding in the active sites of LPMOs as compared to GHs. Hence, the affinity and the specificity of the CBMs may be important for targeting of the LPMOs activity to specific sites at the surface of complex insoluble substrates. Removal of the cellulose binding CBM2 from LPMOs of AA10 enzymes (*Sc*LPMO10C and *Tf*LPMO9B) resulted in a two-fold reduction in cellulolytic activity towards cellulosic substrates *in vitro* [56,57]. By contrast, the deletion of a CBM1 from an AA9 enzyme (*Nc*LPMO9C) did not affect the LPMO activity against cellulosic substrate but decreased it against xyloglucan [58]. Clearly, additional studies are needed to highlight the role and the contribution of CBMs in the mode of substrate binding and function of LPMOs.

Evolutionary conservation of CBM20 from starch-active LPMOs and GHs

Currently, the family AA13 contains only 14 LPMOs in the CAZy database (http://cazy.org/AA13.html), which only displays entries derived from finished GenBank entries. A BLAST search using *An*AA13 as a query uncovered a total of 75 LPMO homologues possessing CBM20 modules. The *A. terreus* LPMO (*At*LPMO13A), which is currently not in the CAZy database, is a close homologue to the characterized *An*AA13, sharing 75% identity, and have confirmed in this study the binding functionality of this enzyme as starch specific LPMO. An analysis of all the AA13 family enzymes in CAZy database showed that the catalytic modules share 60−80% sequence identity encompassing the conserved catalytic residues reported for other LPMOs (Fig. S3) [59]. Strikingly, approximately 60% of these sequences possess a C-terminal CBM20, normally associated with amylolytic enzymes [22], which corresponds to a twice as high relative occurrence of CBMs as compared to average LPMOs. This suggests an evolutionary advantage conferred by the CBM20 in the breakdown of starch, which applies for both LPMOs and GHs.

The alignment of sequences showed that CBM20 appended to AA13 are highly similar to counterparts appended to GH13 and GH15 (Fig. 4 and S4) [60]. Overall, the three out of four consensus starch-binding site (SBS) 1 and 2 carbohydrate-binding residues (corresponding to W543, K578, W590 and W563 of *A. niger* glucoamylase [53]) are

conserved in CBM20s from fungal AA13, GH13 and GH15. It has been shown that substitutions of these residues cause substantial affinity losses for starchy substrates [23,54]. At SBS1 one of tryptophans (W590 in *A. niger* glucoamylase) is substituted by tyrosine in several AA13 proteins, including *Mo*LPMO13A, and glycine in one AA13 protein. The change from a tryptophan to a tyrosine is subtle and likely does not cause major affinity changes. Furthermore, two other polar residues at SBS1 (corresponding to K578 and N595 of *A. niger* glucoamylase) that are known to form hydrogen bonds with ligands such as starch, were conserved in all except the AA13 protein form *Verticillium dahliae*. The conservation of functional residues in LPMO-associated CBM20s is in agreement with the ITC affinity measurements that show comparable binding affinity and thermodynamics to fungal CBM20s. To investigate if CBM20 from LPMOs could be distinguished from counterparts in hydrolases *e.g.* α-amylases and glucoamylases, we performed a phylogenetic analysis of all CBM20 sequences from LPMOs and 24 amylolytic hydrolases (Fig. S5). Notably, the CBM20 sequences clustered largely based on taxonomy, with no apparent regard to nature of the cognate catalytic module (*i.e*. LPMO or hydrolase). These finding suggest a strong pressure to retain the CBM20 as a common binding scaffold during the evolution of catalytically diverse oxidative LPMOs and hydrolases targeting starch.

In conclusion, this study presents the first quantitative data on the binding of modular starch active LPMOs possessing a CBM20 starch binding module. The measured binding affinity and thermodynamic fingerprint revealed a similar binding pattern to canonical CBM20-containing amylolytic enzymes. The binding of the enzyme is thus likely mediated by the CBM20, but the presence of lower affinity binding sites on the catalytic module cannot be excluded. The glycosylation of the enzymes seems to hamper binding as judged by the lower binding stoichiometry obtained from the ITC data, so further work on non-glycosylated enzymes, preferably from fungal hosts, is needed to get additional insight into the affinity of the catalytic modules of LPMOs of AA13 to starch and the contribution of the CBM20 to activity. Analysis of functional residue conservation indicates that the binding function of CBM20 is highly conserved in LPMOs, which is also supported by phylogenetic analysis showing that CBM20 from LPMOs and hydrolases are indistinguishable. The conservation of the CBM20 binding scaffold in the evolution of different oxidative and hydrolytic starch-degrading enzymes asserts the pivotal role of these modules in targeting granular starch.

Acknowledgements

This work was supported by the Novo Nordisk foundation with a “Biotechnology-based Synthesis and Production” grant (NNF12OC0000769). Carlsberg foundation is acknowledged for instrument grants for purchase of ITC and DSC microcalorimeters. A. H. Viborg is thanked for his CAZy research tools.

References

1. Alvani K, Qi X, Tester RF. Gelatinisation properties of native and annealed potato starches. *Starch/Staerke*. 2012;64(4):297-303.
2. Zeeman SC, Kossmann J, Smith AM. Starch: its metabolism, evolution, and biotechnological modification in plants. *Annu Rev Plant Biol*. 2010;61:209-234.
3. Sajilata MG, Singhal RS, Kulkarni PR. Resistant Starch: A review. *Compr Rev Food Sci Food Saf*. 2006;5(1):1-17.
4. Buléon A, Colonna P. Physicochemical behaviour of starch in food applications. In: *The Chemical Physics of Food*.; 2007:20-67.
5. Singh N, Singh J, Kaur L, Sodhi NS, Gill BS. Morphological, thermal and rheological properties of starches from different botanical sources. *Food Chem*. 2003;81:219-231.
6. Pérez S, Bertoft E. The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review. *Starch/Staerke*. 2010;62(8):389-420.
7. Gallant DJ, Bouchet B, Baldwin PM. Microscopy of starch: evidence of a new level of granule organization. *Carbohydr Polym*. 1997;32(3-4):177-191.
8. Buléon A, Colonna P, Planchot V, Ball S. Starch granules: structure and biosynthesis. *Int J Biol Macromol*. 1998;23(2):85-112.
9. Tester RF, Karkalas J, Qi X. Starch - composition, fine structure and architecture. *J Cereal Sci*. 2004;39(2):151-165.
10. Tester RF, Qi X, Karkalas J. Hydrolysis of native starches with amylases. *Anim Feed Sci Technol*. 2006;130(1-2):39-54.
11. Jacobs H, Delcour JA. Hydrothermal modifications of granular starch, with retention of the granular structure: A review. *J Agric Food Chem*. 1998;46(8):2895-2905.
12. Spier F, Zavareze EDR, Marques e Silva R, Elias MC, Dias ARG. Effect of alkali and oxidative treatments on the physicochemical, pasting, thermal and morphological properties of corn starch. *J Sci Food Agric*. 2013;93(9):2331-2337.
13. van der Maarel MJEC, van der Veen B, Uitdehaag JCM, Leemhuis H, Dijkhuizen L. Properties and applications of starch-converting enzymes of the α-amylase family. *J Biotechnol*. 2002;94(2):137-155.
14. Sun H, Zhao P, Ge X, Xia Y, Hao Z, Liu J, Peng M. Recent advances in microbial raw starch degrading enzymes. *Appl Biochem Biotechnol*. 2010;160(4):988-1003.
15. Robertson GH, Wong DWS, Lee CC, Wagschal K, Smith MR, Orts WJ. Native or raw

starch digestion: A key step in energy efficient biorefining of grain. *J Agric Food Chem*. 2006;54(2):353-365.

16. Gilbert HJ, Knox JP, Boraston AB. Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr Opin Struct Biol*. 2013;23(5):669-677.
17. Guillén D, Sánchez S, Rodríguez-Sanoja R. Carbohydrate-binding domains: multiplicity of biological roles. *Appl Microbiol Biotechnol*. 2010;85(5):1241-1249.
18. Hervé C, Rogowski A, Blake AW, Marcus SE, Gilbert HJ, Knox JP. Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *Proc Natl Acad Sci*. 2010;107(34):15293-15298.
19. Cuskin F, Flint JE, Gloster TM, Morland C, Baslé A, Henrissat B, Coutinho PM, Strazzulli A, Solovyova AS, Davies GJ, et al. How nature can exploit nonspecific catalytic and carbohydrate binding modules to create enzymatic specificity. *Proc Natl Acad Sci U S A*. 2012;109(51):20889-20894.
20. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2014;42(D1):490-495.
21. Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J*. 2004;382:769-781.
22. Machovic M, Janecek S. Starch-binding domains in the post-genome era. *Cell Mol Life Sci*. 2006;63(23):2710-2724.
23. Janecek S, Svensson B, MacGregor EA. Structural and evolutionary aspects of two families of non-catalytic domains present in starch and glycogen binding proteins from microbes, plants and animals. *Enzyme Microb Technol*. 2011;49(5):429-440.
24. Juge N, Nøhr J, Le Gal-Coëffet M-F, Kramhøft B, Furniss CSM, Planchot V, Archer DB, Williamson G, Svensson B. The activity of barley α-amylase on starch granules is enhanced by fusion of a starch binding domain from *Aspergillus niger* glucoamylase. *Biochim Biophys Acta*. 2006;1764(2):275-284.
25. Chen LJ, Ford C, Kusnadi A, Nikolov ZL. Improved adsorption to starch of a β-galactosidase fusion protein containing the starch-binding domain from *Aspergillus* glucoamylase. *Biotechnol Prog*. 1991;7(3):225-229.
26. Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JCN, Johansen KS, Krogh KBRM, Jørgensen CI, Tovborg M, Anthonsen A, et al. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci*. 2011;108(37):15079-15084.

27. Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunæs AC, Stenstrøm Y, Mackenzie A, Sørlie M, Horn SJ, Eijsink VGH. Cleavage of cellulose by a CBM33 protein. *Protein Sci*. 2011;20(9):1479-1483.
28. Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WGT, Ludwig R, Horn SJ, Eijsink VGH, Westereng B. Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci*. 2014;111(17):6287-6292.
29. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SWA, Koetsier MJ, van Berkel WJH, Gruppen H, Kabel MA. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels*. 2015;8(1):101.
30. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørlie M, Eijsink VGH. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*. 2010;330(6001):219-222.
31. Crouch LI, Labourel A, Walton PH, Davies GJ, Gilbert HJ. The contribution of non-catalytic carbohydrate binding modules to the activity lytic polysaccharide monooxygenases. *J Biol Chem*. 2016:1-15.
32. Lo Leggio L, Simmons TJ, Poulsen JCN, Frandsen KEH, Hemsworth GR, Stringer MA, von Freiesleben P, Tovborg M, Johansen KS, De Maria L, et al. Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. *Nat Commun*. 2015;6:5961.
33. Vu VV, Beeson WT, Span EA, Farquhar ER, Marletta MA. A family of starch-active polysaccharide monooxygenases. *Proc Natl Acad Sci*. 2014;111(38):13822-13827.
34. Vester-Christensen MB, Abou Hachem M, Naested H, Svensson B. Secretory expression of functional barley limit dextrinase by *Pichia pastoris* using high cell-density fermentation. *Protein Expr Purif*. 2010;69(1):112-119.
35. Silvanovich MP, Hill RD. Affinity chromatography of cereal α-amylase. *Anal Biochem*. 1976;73(2):430-433.
36. Juge N, Andersen JS, Tull D, Roepstorff P, Svensson B. Overexpression, purification, and characterization of recombinant barley α-amylases 1 and 2 secreted by the methylotrophic yeast *Pichia pastoris*. *Protein Expr Purif*. 1996;8(2):204-214.
37. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. *Proteomics Protoc Handb*. 2005:571-607.
38. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.

Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389-3402.

39. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res*. 2015;43(D):222-226.
40. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM, Buso N, Lopez R. The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res*. 2015;43(W):580-584.
41. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. 2014;42(W):320-324.
42. Huson DH, Scornavacca C. Dendroscope 3 : An interactive tool for rooted phylogenetic trees and networks. *Syst Biol*. 2012;61(6):1061-1067.
43. Dimarogona M, Topakas E, Olsson L, Christakopoulos P. Lignin boosts the cellulase performance of a GH-61 enzyme from *Sporotrichum thermophile*. *Bioresour Technol*. 2012;110:480-487.
44. Koseki T, Mese Y, Fushinobu S, Masaki K, Fujii T, Ito K, Shiono Y, Murayama T, Iefuji H. Biochemical characterization of a glycoside hydrolase family 61 endoglucanase from *Aspergillus kawachii*. *Appl Microbiol Biotechnol*. 2008;77(6):1279-1285.
45. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels*. 2012;5(1):79.
46. Penninga D, van der Veen BA, Knegtel RMA, van Hijum SAFT, Rozeboom HJ, Kalk KH, Dijkstra BW, Dijkhuizen L. The raw starch binding somain of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251. *J Biol Chem*. 1996;271(51):32777-32784.
47. Dean R, Van Kan JAL, Pretorius ZA, Hammond-Kosack KE, Di Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J, et al. The Top 10 fungal pathogens in molecular plant pathology. *Mol Plant Pathol*. 2012;13(4):414-430.
48. Wilson RA, Talbot NJ. Under pressure: investigating the biology of plant infection by *Magnaporthe oryzae*. *Nat Rev Microbiol*. 2009;7(3):185-195.
49. Louis B, Waikhom SD, Roy P, Bhardwaj PK, Singh MW, Chandradev SK, Talukdar NC. Invasion of S*olanum tuberosum* L. by *Aspergillus terreus*: a microscopic and proteomics insight on pathogenicity. *BMC Res Notes*. 2014;7(1):350.
50. Kück U, Bloemendal S, Teichert I. Putting fungi to work: Harvesting a cornucopia of drugs, toxins, and antibiotics. Heitman J, ed. *PLoS Pathog*. 2014;10(3):e1003950.

51. Glaring MA, Baumann MJ, Abou Hachem M, Nakai H, Nakai N, Santelia D, Sigurskjold BW, Zeeman SC, Blennow A, Svensson B. Starch-binding domains in the CBM45 family - low-affinity domains from glucan, water dikinase and α-amylase involved in plastidial starch metabolism. *FEBS J*. 2011;278(7):1175-1185.
52. Sigurskjold BW, Christensen T, Payre N, Cottaz S, Driguez H, Svensson B. Thermodynamics of binding of heterobidentate ligands consisting of spacer-connected acarbose and β-cyclodextrin to the catalytic and starch-binding domains of glucoamylase from *Aspergillus niger* shows that the catalytic and starch-binding sites are in close proximity in space. *Biochemistry*. 1998;37(29):10446-10452.
53. Sorimachi K, Le Gal-Coëffet MF, Williamson G, Archer DB, Williamson MP. Solution structure of the granular starch binding domain of *Aspergillus niger* glucoamylase bound to β-cyclodextrin. *Structure*. 1997;5(5):647-661.
54. Williamson MP, Le Gal-Coëffet MF, Sorimachi K, Furniss CSM, Archer DB, Williamson G. Function of conserved tryptophans in the *Aspergillus niger* glucoamylase 1 starch binding domain. *Biochemistry*. 1997;36(24):7535-7539.
55. Sigurskjold BW, Svensson B, Williamson G, Driguez H. Thermodynamics of ligand-binding to the starch-binding domain of glucoamylase from *Aspergillus niger*. *Eur J Biochem*. 1994;225(1):133-141.
56. Forsberg Z, Mackenzie AK, Sorlie M, Rohr ÅK, Helland R, Arvai AS, Vaaje-Kolstad G, Eijsink VGH. Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. *Proc Natl Acad Sci*. 2014;111(23):8446-8451.
57. Arfi Y, Shamshoum M, Rogachev I, Peleg Y, Bayer EA. Integration of bacterial lytic polysaccharide monooxygenases into designer cellulosomes promotes enhanced cellulose degradation. *Proc Natl Acad Sci*. 2014;111(25):9109-9114.
58. Borisova AS, Isaksen T, Dimarogona M, Kognole AA, Mathiesen G, Várnai A, Røhr ÅK, Payne CM, Sørlie M, Sandgren M, et al. Structural and functional characterization of a lytic polysaccharide monooxygenase with broad substrate specificity. *J Biol Chem*. 2015;290(38):22955-22969.
59. Hemsworth GR, Johnston EM, Davies GJ, Walton PH. Lytic polysaccharide monooxygenases in biomass conversion. *Trends Biotechnol*. 2015;33(12):747-761.
60. Christiansen C, Abou Hachem M, Janecek S, Viksø-Nielsen A, Blennow A, Svensson B. The carbohydrate-binding module family 20 - diversity, structure, and function. *FEBS J*. 2009;276(18):5006-5029.

Figure legends

Fig. 1. SDS-PAGE showing Mark 12 protein standard (Invitrogen) in the left lane and recombinant purified *Mo*LPMO13A and *At*LPMO13A in lanes A and B, respectively.

Fig. 2. *Mo*LPMO13A binding to insoluble wheat starch. The binding of 2.3 μM (black circles) and 4.7 μM (grey circles) enzyme to starch (1−45 mg/ml) was performed (see Materials and Methods) and the fraction of bound *Mo*LPMO13A versus wheat starch concentration is plotted. The fits of a one site binding model to the data are shown (grey lines). The means of duplicates ± standard deviations are shown.

Fig. 3. Isothermal titration calorimetry (ITC) of the binding of *At*LPMO13A to β-cyclodextrin. Binding thermogram (top), and normalized integrated heat response data (lower panel, black squares) are shown and the fit (red line) of a one site binding model to the data. Titrations were performed in 10 mM MES, pH 6.5 at 25 °C.

Fig. 4. Sequence alignment of CBM20 domains from *At*LPMO13A, *Mo*LPMO13A and fungal characterized GH13 α-amylases and GH15 glucoamylases. Fully conserved residues are in white letters and red background, the SBS1 residues (two tryptophans and lysine) are marked with red stars, and SBS2 tryptophan is marked with black star. The UniProt accessions and organisms are indicated.

Figures

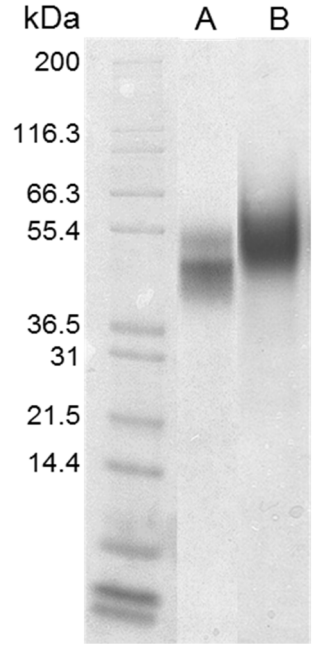


Figure 1.

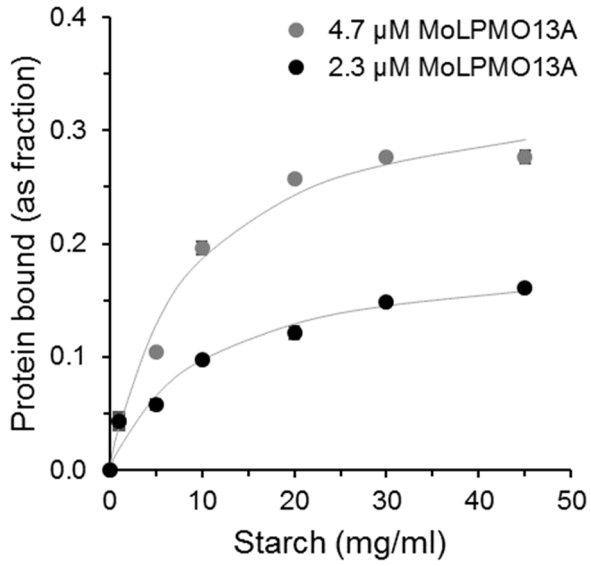


Figure 2.

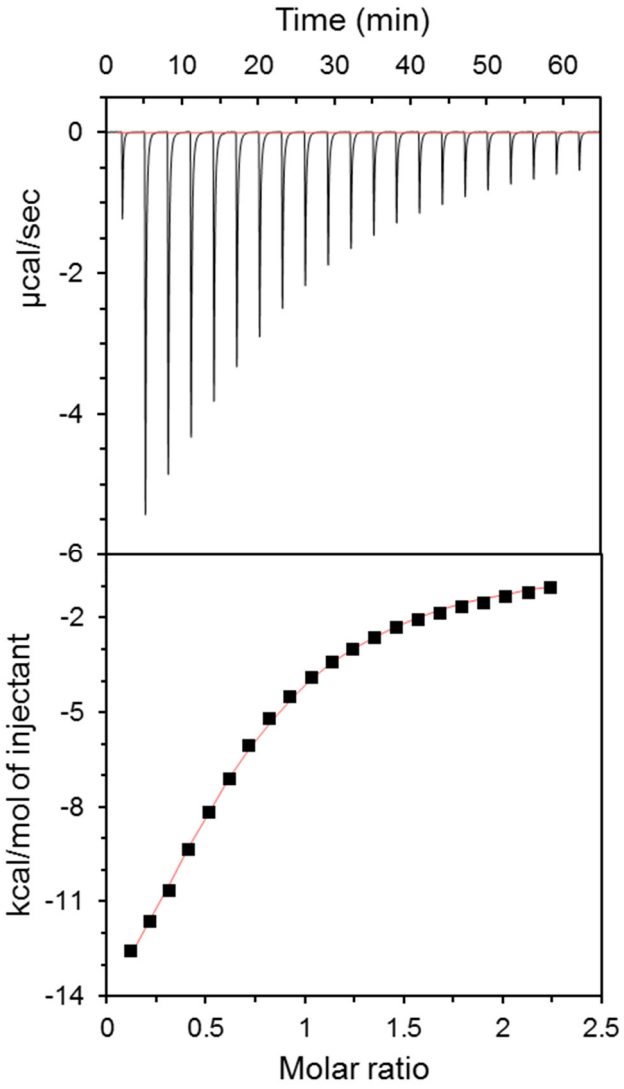


Figure 3.

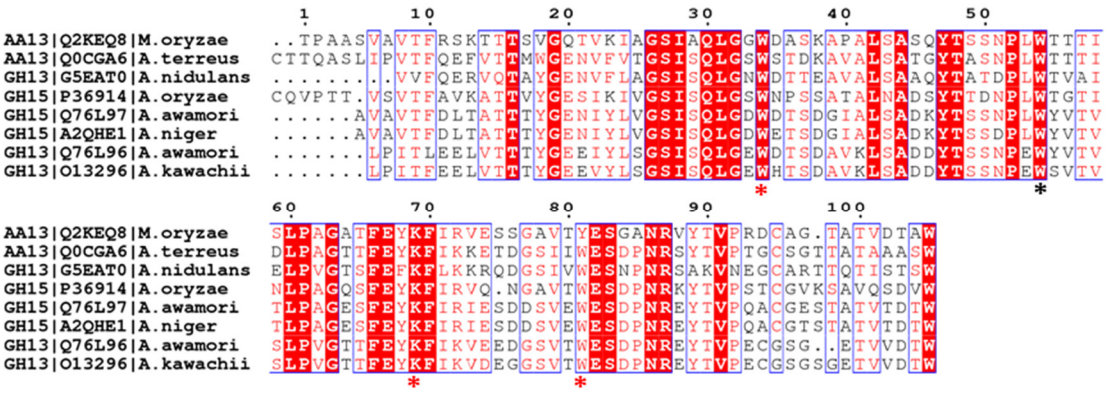


Figure 4.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Fig. S1. Predicted *N*- and *O*-glycosylation sites in *At*LPMO13A and *Mo*LPMO13A (see materials and methods). Asparagines predicted to be *N*-glycosylated are coloured red (no *N*-glycosylation is predicted in *Mo*LPMO13A). Predicted *O*-glycosylated residues are coloured blue. Catalytic modules are coloured grey and CBM20 – bright green.

Fig. S2. Differential scanning calorimetry (DSC) thermograms representing a two-state protein unfolding curves for *At*LPMO13A and *Mo*LPMO13A.

Fig. S3. Sequence alignment of the catalytic domains of all AA13 LPMOs from the CAZy database. Fully conserved residues are shown in white letters on red background, the copper coordinating histidines are marked with red stars, and other residues (G, Q, Y) contributing in active site marked with black stars. GenBank accessions and organisms are shown.

Fig. S4. Alignment of CBM20 modules from AA13 LPMOs and fungal GH13 α-amylases and GH15 glucoamylases. Fully conserved residues are in white letters on red background, the SBS1 two tryptophans and lysine are marked (red stars), and the SBS2 tryptophan is marked (black star). Other suggested SBS1 and SBS2 are marked (blue dots). *At*LPMO13A and *Mo*LPMP13A sequences are highlighted in light blue colour. Amylolytic enzymes are highlighted in grey colour. Accession numbers and organisms are shown.

Fig. S5. Phylogram of CBM20 modules from AA13 LPMOs and fungal GH13 α-amylases and GH15 glucoamylases. Characterized AA13s are highlighted in pink colour. *At*LPMO13A and *Mo*LPMP13A are highlighted in light blue colour. AA13 enzymes are marked as empty circles, GH13s – empty squares, and GH15s – black circles.

Supporting information

>AtLPMO13A

MLLTVLAVVGCFTAVNGHGYLTIPASRTRLGFETGIDTCPECSILEPVTAWPDLEAAQVG
RSGPCGYNARVSVDYNQPSEYWGNEPVVTYTSGEVVEVQWCVDANGDHGGMFTYGICQNQ
TLVDKFLTPGYLPTNEEKQAAEDCFLDGELKCKDVSGQTCGYNPDCTEGAACWRNDWFTC
NAFQANTARACQGVDGASLNSCKTTIAGGYTVTKRIKIPDYSSDHTLLRFRWNSFQTAQV
YLHCADIAIAGSGGGTTSKSTTSTTSTTSTSRSTSTSAPTTTSSASTATPICTTQASLIP
VTFQEFVTTMWGENVFVTGSISQLGSWSTDKAVALSATGYTASNPLWTTTIDLPAGTTFE
YKFIKKETDGSIIWESDPNRSYTVPTGCSGTTATAAASWR

>MoLPMO13A

MKWSVIQALALASGVQAHGYLTFPMSRTGLNAQAGPDTCPECTILEPVTAWPDLDSAQVG
RSGPCGYNARVSVDYNQPGPRWGSAPVVTYKGGDVADVQWCVDNNGDHGGMFTYRICQDQ
ALVDKLLTPGYLPSEAEKQAAENCFRAGTLPCTDVNGQSCGYSPDCSPGQACWRNDWFTC
KGFQDTKCRGVDNAPLNSCYTSIAGGYTTPQIYLTCADIKITAPDSQSPPTTTTTSTPAS
PPPTSCATPAASVAVTFRSKTTTSVGQTVKIAGSIAQLGGWDASKAPALSASQYTSSNPL
WTTTTISLPAGATFEYKFIRVESSGAVTYESGANRVYTVPRDCAGTATVDTAWK

Figure S1.

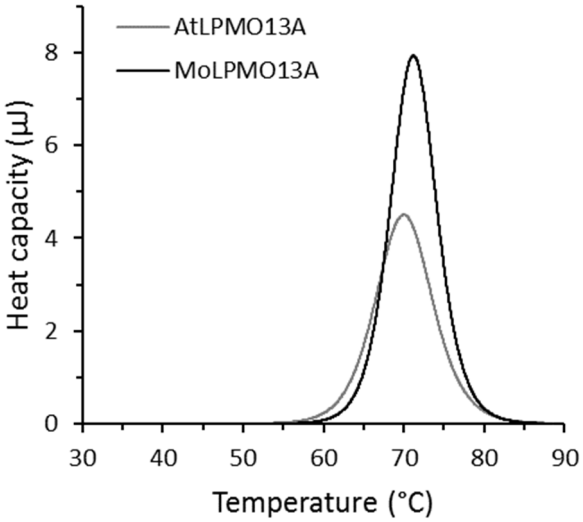


Figure S2.

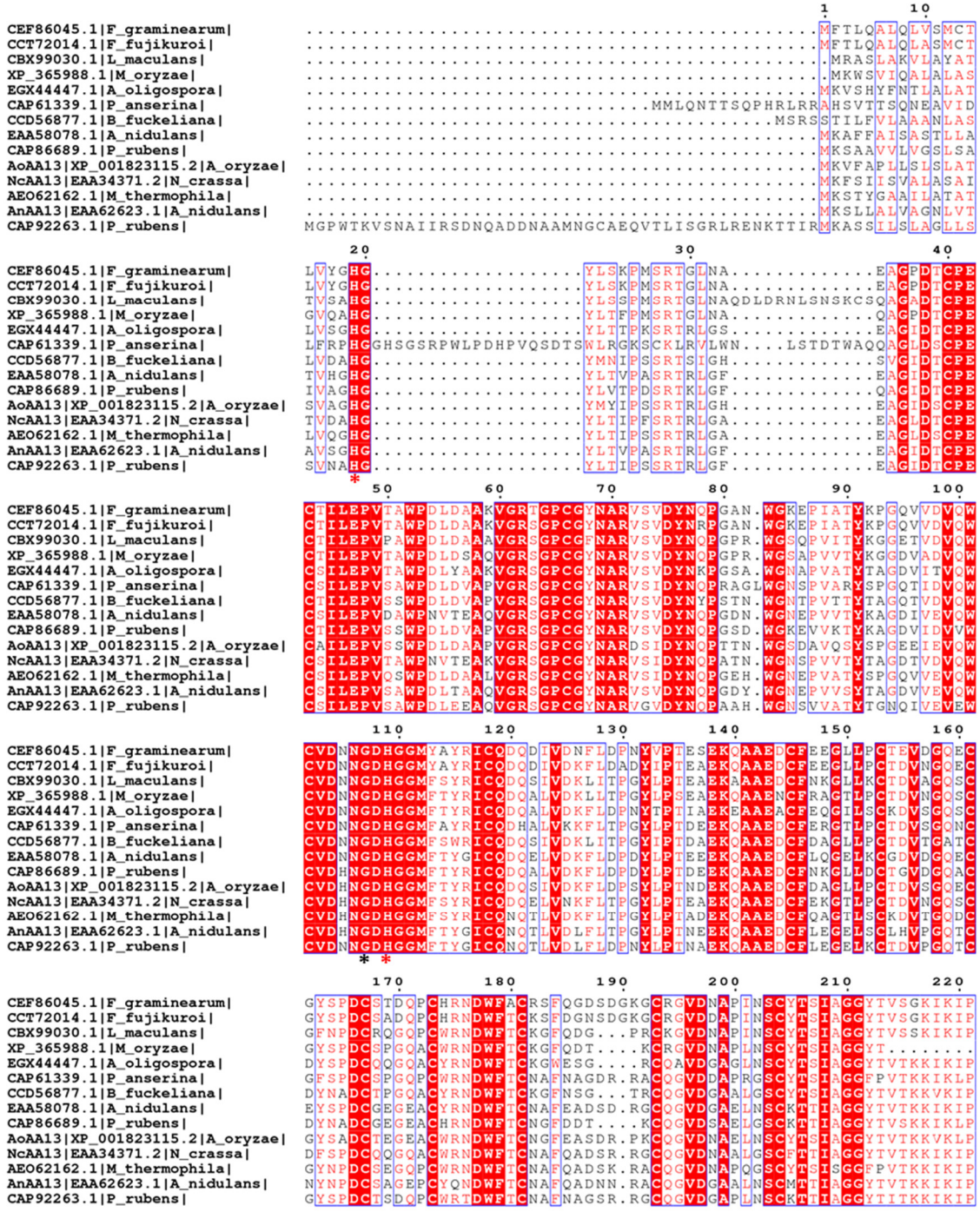

CEF86045.1|F_graminearum|
CCT72014.1|F_fujikuroi|
CBX99030.1|L_maculans|
XP_365988.1|M_oryzae|
EGX44447.1|A_oligospora|
CAP61339.1|P_anserina|
CCD56877.1|B_fuckeliana|
EAA58078.1|A_nidulans|
CAP86689.1|P_rubens|
AoAA13|XP_001823115.2|A_oryzae|
NcAA13|EAA34371.2|N_crassa|
AEO62162.1|M_thermophila|
AnAA13|EAA62623.1|A_nidulans|
CAP92263.1|P_rubens|

Figure S3.

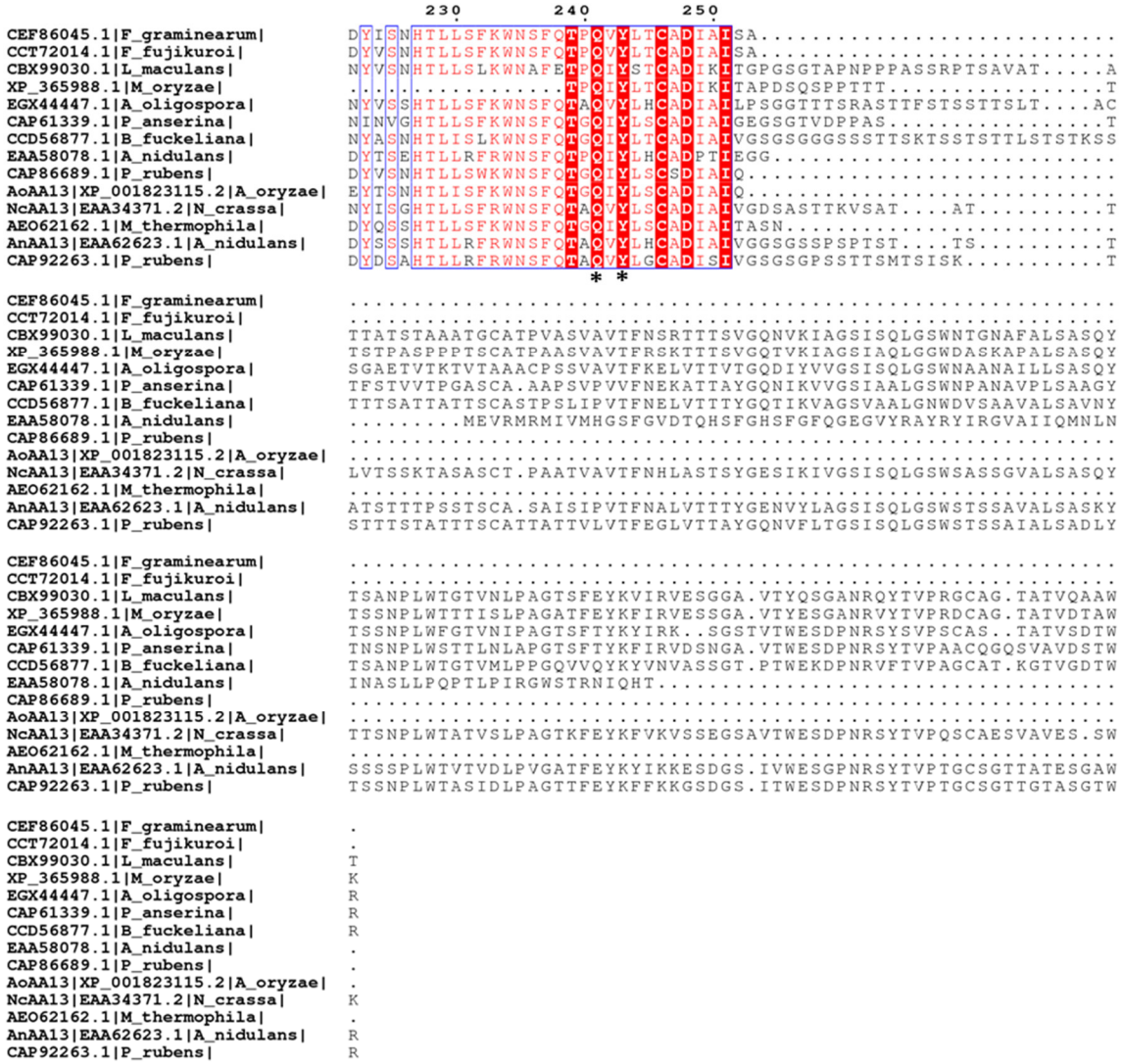


Figure S3. Continued

Figure S4.

Figure S4. Continued

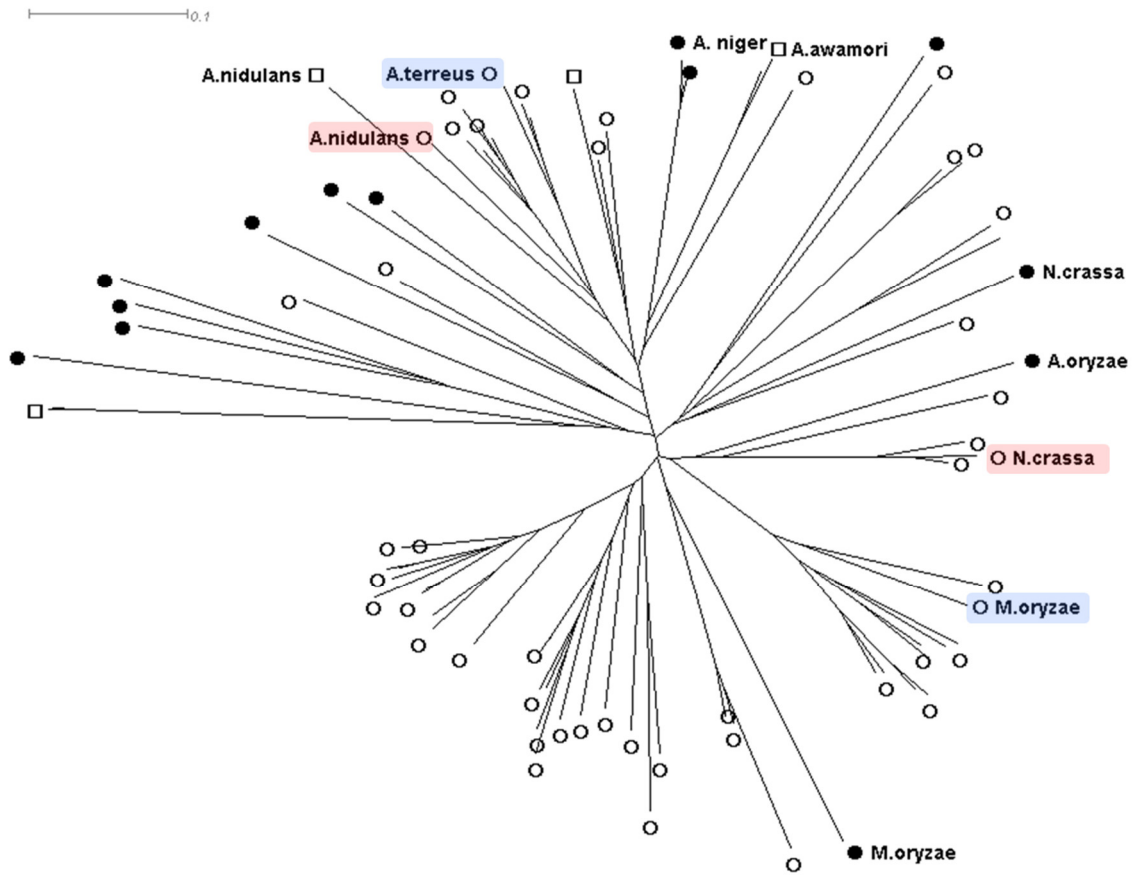

0.1
A. niger
A.awamori
A.nidulans
A.terreus
A.nidulans
N.crassa
A.oryzae
N.crassa
M.oryzae
M.oryzae

Figure S5.

CHAPTER 4. PAPER III

The *Fusarium graminearum* lytic polysaccharide monooxygenase *Fg*LPMO9A cleaves xyloglucan independent on the backbone substitution pattern

The *Fusarium graminearum* lytic polysaccharide monooxygenase *Fg*LPMO9A cleaves xyloglucan independent on the backbone substitution pattern

Laura Nekiunaite[1], Dejan Petrovic[2], Bjørge Westereng[2], Gustav Vaaje-Kolstad[2], Maher Abou Hachem[1], Aniko Varnai[2], Vincent G.H. Eijsink[2*]

1 Enzyme and Protein Chemistry, Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark

2 Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Aas, Norway

***** Corresponding author. Email address: vincent.eijsink@nmbu.no.

Abstract

The discovery of lytic polysaccharide monooxygenases (LPMOs) has drastically changed our current perception of enzymatic biomass conversion, both in industrial settings and in biology. The genomes of plant pathogenic fungi typically contain several LPMO genes that encode enzymes potentially important for plant infection and/or useful as biocatalysts. We describe the cloning, overexpression and characterized an LPMO from *Fusarium graminearum* called *Fg*LPMO9A. *Fg*LPMO9A showed a hitherto not observed substrate profile. The enzyme catalyzes mixed C1/C4 oxidative cleavage of cellulose and xyloglucan, but, in contrast to other xyloglucan active LPMOs, it is not active towards other (1,4)-linked β-glucans nor towards cellohexaose. Furthermore, *Fg*LPMO9A possesses unprecedented broad specificity when acting on xyloglucan, where it can cleave any glycosidic bond in the β-glucan main chain, regardless of the presence of xylose substitutions. Interestingly, xyloglucan activity was higher at 15 °C than at 45 °C, suggesting that substrate rigidity contributes to enzyme activity. This temperature effect was much weaker in experiments with xyloglucan coated cellulose. Interestingly, when incubated with xyloglucan-coated cellulose, *Fg*LPMO9A did only release xyloglucan products, whereas cellulose conversion was inhibited, suggesting that removal of hemicellulose maybe a true and useful function of this LPMO during cellulose conversion.

Keywords: AA9; cellulose; *Fusarium graminearum*; lytic polysaccharide monooxygenase; xyloglucan**.**

Abbreviations: AA, auxiliary activities; DP, degree of polymerization; GH, glycoside hydrolase; LPMO, lytic polysaccharide monooxygenase; MALDI-ToF MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry.

INTRODUCTION

Plants have evolved multiple defense mechanisms against microbial pathogens, varying from defense systems triggered by recognition of microbe-associated molecular patterns, to physical barriers, such as the plant cell wall [1,2]. The plant cell wall is the major structural defense barrier against fungal and bacterial pathogens. The cell wall composite consists of cellulose microfibrils embedded in a matrix of hemicellulose and lignin, with cross-linking between the polymers and a multi-layer structure. The primary cell wall, i.e. the thin outer layer of plant cell walls, also contains pectins and structural proteins, which are much less abundant in the secondary walls [3–5]. Plant pathogenic fungi have to overcome this complex plant cell wall barrier to attack and colonize the host tissue [6]. Production of enzymes that deconstruct host cell wall polymers into mono- and oligosaccharides not only grants access to the plant tissue but also provides fuel for growth and reproduction [7]. *Fusarium graminearum* is a highly destructive pathogen of all cereal species and causes Fusarium head blight disease on wheat and barley. In addition, this devastating fungus produces mycotoxins, which pose a threat to human and animal health [8]. The genome of *F. graminearum* contains almost 600 predicted carbohydrate-active enzymes (CAZymes). These CAZymes include a variety of glycoside hydrolases (GHs) as well as 18 putative lytic polysaccharide monooxygenases (LPMOs) [9]. *F. graminearum*, an efficient degrader of plant biomass, secretes cell wall degrading enzymes, including cellulases, xylanases and pectinases as well as a broad range of accessory enzymes, including LPMOs [10–13].

In contrast to typical glycoside hydrolases (GHs), LPMOs are copper-dependent redox enzymes that oxidatively cleave polysaccharides using molecular oxygen and an electron donor [14–17]. These enzymes are classified as auxiliary activities (AAs) in the Carbohydrate-Active enZyme database, comprising four families, AA9, AA10, AA11, and AA13 (CAZy; http://cazy.org) [18]. LPMOs cleave polysaccharides while oxidizing one of the new chain ends at the C1 or C4 position, thus contributing to substrate depolymerization while increasing accessibility of the substrate to conventional GHs [14,17,19,20]. Since their discovery in 2010 [14], LPMOs with preference for various plant polysaccharides, such as cellulose [21,22], xyloglucan and other (1,4)-linked β-glucans [23], starch [24] and xylan [25] have been identified. LPMOs have become a very important

ingredients in commercial enzyme cocktails for industrial biomass conversion, such as Cellic CTec2 [26,27], owing to their synergy with GHs in improving saccharification yields [28–30].

In the present study, we cloned and characterized an AA9 LPMO from *F. graminearum*, hereafter referred to as *Fg*LPMO9A. The N-terminal AA9 domain in *Fg*LPMO9A is followed by approximately a 100 residues C-terminal domain for which there is no functional prediction. We studied the action of *Fg*LPMO9A on various plant polysaccharides including cellulose and xyloglucan. This is the first *F. graminearum* LPMO to be characterized and the enzyme turned out to cleave β-1,4-linked glycans, performing C1/C4 mixed oxidation. Unlike other LPMOs described so far, substitution on the β-1,4-glycan backbone by xylosyl units at the C6 position, as in xyloglucan did not affect enzyme action. Thus, *Fg*LPMO9A possesses unprecedented broad specificity when acting on xyloglucan.

RESULTS

Heterologous expression of FgLPMO9A from F. graminearum

Recombinant *Fg*LPMO9A was successfully expressed by growing *P. pastoris* in BMGY medium for 48 hours. The enzyme was purified in two chromatographic steps to ≈95% purity, as confirmed by SDS-PAGE (Addition file, Fig. S1). The protein band representing *Fg*LPMO9A appears as a slightly smeared band in the 60–70 kDa range, whereas the theoretical molecular mass calculated from the amino acid sequence is 32 kDa. The observed mass difference is likely to due to *O*- and/or *N*-glycosylation, which are predicted for residues Ser/Thr (234, 235, 238, 240, 241, 243, 245, 247, 249, 251, 257, 262, 287, 292) and Asn211, respectively. Attempts to deglycosylate the protein by endoglycosidases removing *N*-glycosylations did not alter its electrophoretic mobility (data not shown). Similar observations have been made for other LPMOs expressed in yeast [31–33].

Substrate specificity

In order to determine substrate specificity, *Fg*LPMO9A was incubated with a wide range of oligo- and polysaccharides in the presence of an electron donor. Studies with insoluble substrates showed that the enzyme is capable of cleaving cellulose (Fig. 1), but not chitin (results not shown). No activity was detected towards shorter cellooligosaccharides ((Glc_{3-6}); results not shown). Among the tested hemicelluloses, LPMO activity was only observed towards tamarind xyloglucan and xyloglucan oligosaccharides (see below for details). No activity was observed towards the xyloglucan-heptamer (XXXG), birchwood xylan, wheat arabinoxylan, konjac glucomannan, ivory nut mannan, β-glucan from barley, lichenan from Icelandic moss, starch, and spruce (data not shown).

Activity on amorphous cellulose (PASC)

When using PASC as a substrate, HPAEC analysis of the products generated by *Fg*LPMO9A showed native as well as C1- and C4-oxidized cellooligosaccharides (Fig. 1A). In the absence of an electron donor, no LPMO activity was detected. Notably, it was recently discovered that during HPAEC, C4-oxidized cello-oligomers undergo on-column degradation, yielding native cello-oligomers that lack one glucose unit relative to the original C4-oxidized product [34]. Oxidative cleavage of cellulose and the C1/C4 mixed oxidation pattern were confirmed by MALDI-ToF MS analysis of the released products (Fig. 1B–C). Notably, both types of oxidized products may occur in the non-hydrated and the hydrated form, the latter being a gemdiol for C4-oxidized products and a charged aldonic acid for C1-oxidized products. To further validate product identities, samples were saturated with sodium, thus suppressing the formation of potassium adducts (Fig. 1C). The latter are problematic because the mass difference between potassium and sodium equals the mass of oxygen.

The appearance of native oligosaccharides, such as the sodium adduct of cellohexaose with *m/z* 1013.8 (Fig. 1C), indicates C4 oxidation, as this is commonly observed for C4 oxidizing LPMOs [19]. It has recently been shown that these often observed remarkably high levels of native products are caused by C4 oxidized cellooligomers undergoing on-column degradation that yield native products [34]. The *m/z* 1051.8 signal demonstrates C1 oxidation, since this can only be the sodium-adduct of the sodium salt of the aldonic acid. Oxidation at the C4 position elicits a large signal at (*m/z* 1011.8, for DP 6) because for the ketoaldose, the non-hydrated form is favored, while this is opposite for the lactone generate by C1 oxidation [35]. At neutral conditions C1-oxidized products preferentially appear in the hydrated aldonic acid form (*m/z* 1029.8). After sodium doping most of the signal at *m/z* 1027.8 disappeared indicating that this primarily is a potassium adduct of a non-hydrated oxidized species. However, a small signal remained which likely is the single sodium adduct of a double oxidized, singly hydrated species (4-ketoaldose and aldonic acid). All clusters representing cellooligosacchrides ranging in DP from 4 to 10 (Fig. 1B) had a similar adduct distribution as shown for the DP 6 cluster in Fig. 1C.

Activity on tamarind xyloglucan

In addition to cellulose, *Fg*LPMO9A was also capable of cleaving tamarind xyloglucan (TXG) (Fig. 2 and 3). HPAEC analysis revealed a more complex product profile (Fig. 2) compared to the product profile generated by *Nc*LPMO9C (Fig. S2 in [23]). *Fg*LPMO9A released a mixture of oligosaccharides, some of which corresponded to the native xyloglucan heptasaccharide (XXXG), eluting at 26 min.

Some species eluted before XXXG, which are likely to be shorter native fragments, such as XXX, eluting at 22 min (see Agger et al., 2014). The elution profile in the 50-55 min range was similar to the profile previously published for *Nc*LPMO9C [23]. In addition, a broad range of product species appeared at 26-50 min that have not been observed before, not for *Nc*LPMO9C, nor for the only other xyloglucan-active LPMO described so far, *Pa*LPMO9H [23,36]. Although some of these peaks may correspond to longer native XG oligomers (see black line in Fig. 2), most of the products released from TXG remain unidentified due to the heterogeneity of the substrate and lack of standards.

To confirm oxidative cleavage of TXG by *Fg*LPMO9A, we analyzed the reaction products with MALDI-ToF/MS, which revealed the formation of a broad range of native and oxidized xyloglucan-oligosaccharides (Fig. 3A). Comparison of the products generated by *Fg*LPMO9A to those generated by *Nc*LPMO9C (Fig. 3B) confirmed that *Fg*LPMO9A released a broader range of products than *Nc*LPMO9C. Products generated by both LPMOs included clusters of oxidized products. The first cluster consists of $Hex_nPen_3^{ox}$ oxidized products with *m/z* 1083.7 (n=4), *m/z* 1245.5 (n=5), and *m/z* 1407.5 (n=6). The second cluster consists of $Hex_nPen_6^{ox}$ oxidized products with *m/z* 2452.1 (n=10), *m/z* 2614.1 (n=11), and *m/z* 2776.2 (n=12), where Hex stands for a hexose (glucose or galactose, 162.14 Da) and Pen for a pentose (xylose, 132.11 Da). In addition, *Fg*LPMO9A released unique products that were not detected in reactions with *Nc*LPMO9C, like $Hex_8Pen_4^{ox}$ (*m/z* 1863.7) and $Hex_9Pen_5^{ox}$ (*m/z* 2157.9). As illustrated in Fig. 3C, the fact that all TXG fragments generated by *Nc*LPMO9C comprise a multitude of 3 pentose units indicates that cleavage only has occurred between repeating units of TXG. This assignment is based on the known structure of TXG, in which every fourth glucose unit in the glucan backbone is unsubstituted [37]. Indeed, Agger et al. have previously shown that *Nc*LPMO9C only cleaves at the unsubstituted glucose unit [23]. The fact that *Fg*LPMO9A generates fragments containing a number of pentose units that does not equal a multiple of 3 shows that *Fg*LPMO9A can cleave TXG between two substituted units. This leads to a wide range of products being formed as also shown by the HPAEC chromatograms in Fig. 2.

The clear difference between *Nc*LPMO9C and *Fg*LPMO9A is further illustrated in Fig. 3D. The enzyme simultaneously produces native (*m/z* 1865.7) and oxidized species (*m/z* 1863.7 and 1881.7) which likely reflects mixed (C1/C4) oxidation. Notably, the signal at *m/z* 1879.7 could partly reflect a double-oxidized species.

Most remarkably, the activity of *Fg*LPMO9A towards TXG was considerably higher at 15 °C compared to 45 °C (Fig. 2).

Activity on a complex substrate, xyloglucan-coated PASC

It has previously been shown that certain LPMOs can cleave xylan, only if the xylan is complexed with cellulose [25]. In order to gain further insight into the potential role of *Fg*LPMO9A in plant cell wall degradation, we coated cellulose (PASC) with different hemicelluloses, birchwood xylan, tamarind xyloglucan, ivory nut mannan or konjac glucomannan, and treated the resulting composite polysaccharides with *Fg*LPMO9A, in the presence or absence of ascorbic acid as electron donor. Coating PASC with birchwood xylan, ivory nut mannan or konjac glucomannan did not promote activity of *Fg*LPMO9A towards these three hemicelluloses (data not shown). *Fg*LPMO9A activity on PASC coated with TXG was tested at two different temperatures: 15 °C and 45 °C. The results (Fig. 2) show two things: firstly, the temperature effect is much reduced; secondly, while TXG is still cleaved when complexed to PASC, the PASC itself is no longer cleaved. In the absence of ascorbic acid, no products were formed from PASC or TXG.

Activity on xyloglucan-oligosaccharide (XG14)

To gain a deeper understanding of the *Fg*LPMO9A activity towards TXG, we studied the degradation products generated by the enzyme when incubated with the reduced form of a relatively pure xyloglucan oligosaccharide with DP14 ($XG14^{OH}$). The sequence of this oligomer is $XXXGXXXG^{OH}$, where G^{OH} represents a D-glucitol, reduced glucose, at the reducing end, leading to an *m/z* shift of +2 Da compared to a native species. The product profile was analyzed using HPAEC and MALDI-ToF/MS (Fig. 4A–B). Only in the presence of ascorbic acid, *Fg*LPMO9A generated a variety of products, including the native XXXG peak eluting at 26 min and several additional compounds eluting before XXXG. These latter compounds also appeared in the chromatograms obtained upon degradation of TXG by *Fg*LPMO9A (Fig. 2) and are most likely shorter native fragments. At the same time, several product species eluting after the substrate, at 44-48 min and at 56-64 min, were visible similar to products generated by *Nc*LPMO9C from this same substrate [23].

Oxidative cleavage of $XG14^{OH}$ by *Fg*LPMO9A was confirmed by analysis of the products with MALDI-ToF MS (Fig. 4B). Single C1-oxidation cleavage of $XG14^{OH}$ generates two products that differ with $\Delta m/z$ ±2 from the native ones: the non-reducing end product carrying the C1-oxidation ($\Delta m/z$=-2) and the reducing end product carrying the reduced glucose at the "reducing" end ($\Delta m/z$=+2). On the other hand, single C4-oxidation cleavage generates two products with masses identical to those of native product ($\Delta m/z$=0): the non-reducing end product is a native oligosaccharide and the reducing end product carries a keto-group from the C4-oxidation ($\Delta m/z$ -2)

and the reduced glucitol at the "reducing end" (*Δm/z* +2) (Fig. 4C). The mass spectrum shows a wide-range of native and oxidized xyloglucan oligosaccharides. The occurrence of products with *m/z* 1543.7 ($Hex_6Pen_4^{ox}$), 1837.9 ($Hex_7Pen_5^{ox}$), and 1999.9 ($Hex_8Pen_5^{ox}$), with *Δm/z*=+2 compared to the mass of the native XG-oligosaccharide products demonstrates C1-oxidation. The many products with masses corresponding to native XG-oligosaccharides, such as at *m/z* 1085.4 (Hex_4Pen_3), 1247.5 (Hex_5Pen_3), and 1571.7 (Hex_7Pen_3), confirm C4-oxidation.

DISCUSSION

The discovery of LPMOs has been a significant breakthrough in understanding how plant biomass is degraded in nature, and has also had major technological implications, as illustrated by the fact that, today, these enzymes are important components of enzyme cocktails for industrial biomass conversion [38]. LPMOs are produced not only by saprophytic fungi but also by plant pathogens, suggesting a putative role of these enzymes in pathogenesis. Indeed, it has been shown that most of the putative LPMOs encoded in the genome of *F. graminearum*, including *Fg*LPMO9A, are upregulated during infection of barley and wheat [39]. Although to date the CAZy database harbors more than 400 fungal LPMOs, only about 20 of these enzymes have been characterized, and these are all from saprophytic fungi. In the present study, we selected an LPMO, *Fg*LPMO9A, from the plant pathogen *F. graminearum*, produced it in *P. pastoris*, and studied its substrate specificity. To our knowledge, *Fg*LPMO9A is the first LPMO from a plant pathogen to be characterized.

*Fg*LPMO9A has mixed oxidative regioselectivity, oxidizing both C1 and C4 in insoluble (cellulose) and soluble (xyloglucan) substrates. The enzyme produces relatively large amounts of native cello- and xyloglucan-oligosaccharides (Fig. 1–3). Native products are commonly observed for mixed activity LPMOs, and could emerge when a single polysaccharide chain is cleaved twice, once with C1 and once with C4 oxidation, with the C1 oxidation happening upstream of the C4 oxidation. However, native products are also commonly observed for strictly C4 oxidizing LPMOs and it has recently been shown that the oxidized products of these LPMOs are converted to native species by chemical processes promoted by the conditions of standard HPAEC analysis [34]. Control reactions, where ascorbic acid was excluded from the reaction, showed no formation of oxidized or native xyloglucan- or cello-oligomers, verifying that the *Fg*LPMO9A preparation was free of contaminating background endo-β-1,4-glucanase activity.

So far, only two LPMOs have been reported to cleave xyloglucan, namely *Nc*LPMO9C from *Neurospora crassa* and *Pa*LPMO9H from *Podospora anserina* [20,36]. Both enzymes cleave

xyloglucan exclusively adjacent to the unsubstituted glucosyl unit, *Nc*LPMO9C only yielding C4-oxidized products, whereas *Pa*LPMO9H yields a mixture of C1- and C4-oxidized products with the latter being dominant [36]. Considering the oxidative mechanism, *Fg*LPMO9A is more similar to *Pa*LPMO9H since it oxidizes both C1 and C4. However, it is clear that *Fg*LPMO9A is fundamentally different from both enzymes described above, since *Fg*LPMO9A is not restricted to cleaving the unsubstituted glucosyl, thereby generating a broader spectrum of product species. *Fg*LPMO9A is active on xyloglucan-14-mer (with an octamer backbone), but is not active on the shorter xyloglucan-heptamer (with a tetramer backbone) nor on cello-oligomers up to DP 6. The lack of activity on e.g. cellohexaose separates *Fg*LPMO9A from the other two xyloglucan-active LPMOs described so far and indicates that the catalytic site of *Fg*LPMO9A needs to accommodate a β-1,4-glucan chain with more than six glucose residues in order for catalysis to happen.

The XG-cleaving properties of *Fg*LPMO9A are not commonly found amongst XG-active enzymes, as the majority of GHs cleaving this polymer are selective for unsubstituted backbone glucose residues. The only GHs know to cleave XG between substituted backbone residues are only found in the GH74 family [40,41]. Thus, our findings provide a new activity in the XG enzymatic toolbox. The *F. graminearum* genome only contains one GH74 enzyme, which has not yet been characterized and thus has an unknown function (the GH74 family also harbors enzymes with endoglucanase and exo-xyloglucanase activity). Thus, it may be that *Fg*LPMO9A plays a crucial role in the XG degradative machinery of the fungus by providing the ability to cleave between substituted backbone residues.

*Fg*LPMO9A did not show the ability to cleave xylan, arabinoxylan, glucomannan, mannan, β-glucan, or lichenan. Although the backbone of glucomannan contains β-1,4-glucose-glucose linkages, which are cleaved by e.g. *Nc*LPMO9C, the stretches of glucose-glucose linkages apparently are too short to be recognized by *Fg*LPMO9A. Similarly, despite the similarity between β-1,4-xylan and β-1,4-glucan structures, this LPMO did not show detectable activity on xylan either. Interestingly, there are no LPMOs reported so far that target both xylan and xyloglucan [23,25,36].

An interesting observation was made when comparing the activity of *Fg*LPMO9A towards xyloglucan at unusually low temperatures. After one-hour incubation, substantially more products were formed at 15 °C compared to 45 °C. This may be related to the greater concentration of dissolved O_2 in liquid at lower temperature (approximately twice as much O_2 can be dissolved at 15 °C compared to 45 °C). *A priori*, perhaps a more likely explanation is that the temperature affects the substrate conformation, as also suggested by the fact that the temperature effect was not

observed in the experiments with PASC-xyloglucan composites discussed below. So far, the temperature dependency of LPMO activity has not been investigated thoroughly, and our current results indicate that this may be a side of the enzyme reaction worth studying. Notably, high activity at low temperature is in agreement with the growth habitat of *Fusarium* species, which thrive at lower temperatures in the 15–25 °C range [42–44].

In nature, xyloglucan and cellulose have been proposed to form a network that has major regulatory and load-bearing functions in the plant primary cell walls [3,45,46]. In order to mimic potential natural substrates, we coated xyloglucan, xylan, mannan or glucomannan on cellulose (PASC) and studied the effect of *Fg*LPMO9A on this substrate composite. *Fg*LPMO9A showed activity only on cellulose coated with xyloglucan. Interestingly, while activity of xyloglucan was maintained, activity of cellulose was strongly reduced, suggesting that one natural function of the LPMO could be to make cellulose accessible by removal of hemicellulose. The strong effect of temperature observed for xyloglucan in solution largely disappeared when the xyloglucan was grafted onto the cellulose. The ability of *Fg*LPMO9A to cleave the xyloglucan backbone independently of the decoration of the backbone glucosyl moieties is intriguing, especially since this activity is observed for both xyloglucan in solution and bound to the cellulose (Fig. 2–3). Molecular dynamics simulations suggest that XG adopts different structures in these environments, i.e. a ribbon-like structure in solution and a flat structure when bound to cellulose (generated by the tight interaction between the cellulose chains and the β-glucan backbone) [47]. It is possible that *Fg*LPMO9A recognizes the flat conformation present in the physiologically relevant composite form bound to cellulose, but which in the absence of cellulose is stabilized at low temperature.

LPMOs are thought to have a big impact on the depolymerization of organic matter in nature [14,48]. LPMOs also seem to play a role in the pathogenicity of bacteria such as *Vibrio cholerae* [49], *Listeria monocytogenes* [50] and *Paenibacillus larvae* [51], but there is no data to date showing that LPMOs play a role in fungal pathogenicity. It is conceivable that *Fg*LPMO9A, with its unique ability to randomly cleave xyloglucan, contributes to pathogenesis of *F. graminearum*. The enzyme could contribute to overcoming the plant cell wall barrier, by removing hemicellulose and by making other polymers, such as cellulose more accessible for enzymatic degradation, ultimately leading to tissue invasion and pathogen dissemination.

METHODS

Cloning and expression of the F. graminearum LPMO

The gene encoding *Fg*LPMO9A (UniProt: I1REU9) including its native signal sequence was codon optimized for *Pichia pastoris* (GenScript, NJ, USA). The synthetic gene was inserted into the pPINK-GAP vector [52], yielding pPINK-GAP_FgLPMO9A. This plasmid was linearized with *Afl*II and transformed into freshly prepared electrocompetent *P. pastoris* PichiaPink™ Strain 4 cells (Invitrogen, CA, USA), following the manufacturer's instructions, using a Bio-Rad Gene Pulser II electroporation system (Bio-Rad Laboratories, CA, USA) at 1.8 kV, 25 µF and 200 Ω. Four colonies were picked and re-streaked on fresh PAD (*Pichia* adenine dropout) plates (Invitrogen). Overnight cultures of the four transformants were screened for protein production in BMGY medium (containing 1% (v/v) glycerol).

The best-producing transformant was pre-grown in 20 ml of BMGY medium (containing 1% (v/v) glycerol) in a 100 ml shake flask at 29 °C and 200 rpm for 16 hours. This pre-culture was used to inoculate 0.5 l BMGY medium (containing 1% (v/v) glycerol) in a 2 l baffled shake flask. After 24 hours of incubation at 29 °C and 200 rpm, the medium was additionally supplemented with 1% (v/v) glycerol. After a total incubation time of 48 hours, the cells were harvested by centrifugation at 7 000 × *g* for 15 min and filtered through a 0.2 µm polyethersulfone membrane (Millipore, MA, USA). The supernatant was dialyzed against 50 mM Bis-Tris buffer (pH 6.5) and concentrated to 50 ml with a VivaFlow 50 tangential crossflow concentrator (MWCO 10 kDa, Sartorius Stedim Biotech, Germany).

Purification and Cu(II) saturation of the enzyme

*Fg*LPMO9A was purified using a two-step purification protocol, starting with hydrophobic interaction chromatography (HIC) followed by size exclusion chromatography (SEC). The concentrated culture supernatant was prepared for HIC by slow addition of 2.03 M (final concentration) ammonium sulfate at 20 °C, followed by centrifugation (15 000 × *g*, 15 min, 4 °C). The sample was loaded onto a 5 ml HiTrap Phenyl FF column (GE Healthcare, Sweden) equilibrated with 50 mM Bis-Tris buffer (pH 6.5), containing 2.03 M ammonium sulfate. Proteins bound to the column were eluted using a 25 ml linear gradient from 2.03 M to 0 M ammonium sulfate in 50 mM Bis-Tris buffer (pH 6.5). Collected fractions were analyzed by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE), and the fractions containing *Fg*LPMO9A were pooled and concentrated using Amicon Ultra centrifugal filters (MWCO 10 kDa, Millipore). The sample was saturated with Cu(II) by incubating the enzyme with an excess of $CuSO_4$ (at 3:1

molar ratio of copper:enzyme) for 30 min at room temperature as described previously [49]. Subsequently, the sample was loaded onto a HiLoad 16/60 Superdex 75 size exclusion column (GE Healthcare) in 50 mM Bis-Tris buffer (pH 6.5), using a flow rate of 1.0 ml/min. Fractions containing pure protein were identified using SDS-PAGE and subsequently pooled and concentrated using Amicon Ultra centrifugal filters (MWCO 10kDa, Millipore). Protein concentrations were determined using the Bradford assay (Bio-Rad).

Lytic polysaccharide monooxygenase *Nc*LPMO9C from *Neurospora crassa* (UniProt:Q7SHI8) was expressed in *P. pastoris* X-33 and purified as described previously [33].

Substrates used for enzyme specificity analysis

The following substrates were used for exploring enzymatic activities of the recombinant *Fg*LPMO9A: phosphoric acid swollen cellulose (PASC) prepared from Avicel as described by Wood [53]; xyloglucan (XG) from tamarind seed; xyloglucan-heptamer (XG7) with known composition XXXG (X stands for glucose linked at C6 to xylose); xyloglucan oligomers (XG-oligomers) derived from tamarind xyloglucan mainly containing XXXGXXXG, with 0–3 galactose substitutions (L stands for X where the C2 of the xylose is linked to a galactose; it is not known which X are galactosylated); a reduced 14-mer ($XG14^{OH}$) mainly containing $XXXGXXXG^{OH}$ and variants containing one or more L instead of X (the reducing end D-glucose is reduced to a D-glucitol). *Fg*LPMO9A was also screened against celloologosaccharides (Glc_{3-6}), birchwood xylan, wheat arabinoxylan, konjac glucomannan, ivort ut mannan, barley β-glucan, lichenan from Icelandic moss, starch, and spruce. Birchwood xylan was purchased from Sigma-Aldrich (MO, USA) and all other substrates were purchased from Megazyme (Ireland).

Deglycosylation

Putatively *N*-glycosylated asparagines on *Fg*LPMO9A were deglycosylated by incubating the enzyme with Endo H and PNGase F (New England Biolabs, MA, USA). In analytical experiments, the enzyme (10 µg) was mixed with 2 µl Endo H and 2 µl GlycoBuffer 3 (10X) and H_2O or with 2 µl PNGase F and 2 µl GlycoBuffer 2 (10X) and H_2O in a 20 µl total reaction volume. Both reactions were incubated at 37 °C for 4 hours and mobility shifts were analyzed by SDS-PAGE. *N*-glycosylation sites were predicted using the servers NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) and GlycoEP (http://www.imtech.res.in/raghava/glycoep/). *O*-glycosylation sites were predicted using the servers NetOGlyc (http://www.cbs.dtu.dk/services/NetOGlyc/) and GlycoEP.

Enzyme activity assays

Reaction mixtures contained 1–4 mg/ml substrate, 2 µM *Fg*LPMO9A or *Nc*LPMO9C and 1 mM ascorbic acid in 20 mM Bis-Tris buffer (pH 6.5) in a total volume of 100 µl in 2 ml Eppendorf tubes. Samples were incubated at 15 or 45 °C in an Eppendorf Thermomixer (Eppendorf AG, Germany) with shaking at 900 rpm. After 1 hour incubation the reaction mixtures were placed on ice and the reactions were immediately stopped by separating soluble and insoluble fractions using a 96-well filtrate plate (Millipore) operated by a Millipore vacuum manifold. Control reactions were performed using identical conditions, but in the absence of ascorbic acid.

Analysis of enzyme products

Products generated by the LPMOs were analyzed using MALDI-ToF/mass spectrometry (MS) and high-performance anion exchange chromatography (HPAEC) as follows. MALDI-ToF/MS analysis was performed with an Ultraflex MALDI-ToF/ToF instrument (Bruker Daltonics, Germany) equipped with a Nitrogen 337 nm laser beam as described by Vaaje-Kolstad et al. [14]. The instrument was operated in positive acquisition mode and controlled by the FlexControl 3.3 software package. The data were processed with mMass software [54]. Sodium saturation was obtained by mixing 5 µl sample with 5 µl 50 mM sodium acetate, followed by 30 min incubation, sample spotting and drying.

HPAEC analysis was performed on a Dionex ICS-5000 system (Dionex, CA, USA) equipped with pulsed amperometric detection (PAD) and a CarboPac PA1 analytical column (2 × 250 mm) with a CarboPac PA1 guard column (2 × 50 mm). The system was operated at a flow rate of 0.25 ml/min and kept at 30 °C while running a 75 min stepwise gradient as described by Agger et al. [23]. The data were collected and analyzed using Chromeleon 7.0 (Dionex).

Acknowledgements

This work was funded in part by the Research Council of Norway through grants 214613 & 243663. We thank Roland Ludwig and colleagues for the clone producing *Nc*LPMO9C. LN was funded by the Novo Nordisk foundation with a “Biotechnology-based Synthesis and Production” grant (NNF12OC0000769).

References

1. Grennan AK. Plant response to bacterial pathogens. Overlap between innate and gene-for-gene defense response. *Plant Physiol*. 2006;142(3):809-811.
2. de León IP, Montesano M. Activation of defense mechanisms against pathogens in mosses and flowering plants. *Int J Mol Sci*. 2013;14(2):3178-3200.
3. Carpita NC, Gibeaut DM. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J*. 1993;3(1):1-30.
4. Cosgrove DJ. Growth of the plant cell wall. *Nat Rev Mol Cell Biol*. 2005;6(11):850-861.
5. Doblin MS, Pettolino F, Bacic A. Plant cell walls: the skeleton of the plant world. *Funct Plant Biol*. 2010;37(5):357-381.
6. Laluk K, Mengiste T. Necrotroph attacks on plants: wanton destruction or covert extortion? *Arabidopsis Book*. 2010;8:e0136.
7. Kubicek CP, Starr TL, Glass NL. Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. *Annu Rev Phytopathol*. 2014;52(1):427-451.
8. Goswami RS, Kistler HC. Heading for disaster: *Fusarium graminearum* on cereal crops. *Mol Plant Pathol*. 2004;5(6):515-525.
9. King R, Urban M, Hammond-Kosack MCU, Hassani-Pak K, Hammond-Kosack KE. The completed genome sequence of the pathogenic ascomycete fungus *Fusarium graminearum*. *BMC Genomics*. 2015;16(1):544.
10. Paper JM, Scott-Craig JS, Adhikari ND, Cuomo CA, Walton JD. Comparative proteomics of extracellular proteins *in vitro* and *in planta* from the pathogenic fungus *Fusarium graminearum*. *Proteomics*. 2007;7(17):3171-3183.
11. Yang F, Jensen JD, Svensson B, Jørgensen HJL, Collinge DB, Finnie C. Secretomics identifies *Fusarium graminearum* proteins involved in the interaction with barley and wheat. *Mol Plant Pathol*. 2012;13(5):445-453.
12. Debeire P, Delalande F, Habrylo O, Jeltsch JM, Van Dorsselaer A, Phalip V. Enzymatic cocktails produced by *Fusarium graminearum* under submerged fermentation using different lignocellulosic biomasses. *FEMS Microbiol Lett*. 2014;355(2):116-123.
13. Lionetti V, Giancaspro A, Fabri E, Giove SL, Reem N, Zabotina OA, Blanco A, Gadaleta A, Bellincampi D. Cell wall traits as potential resources to improve resistance of durum wheat against *Fusarium graminearum*. *BMC Plant Biol*. 2015;15:6.
14. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørlie M, Eijsink VGH. An oxidative

enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*. 2010;330(6001):219-222.

15. Hemsworth GR, Davies GJ, Walton PH. Recent insights into copper-containing lytic polysaccharide mono-oxygenases. *Curr Opin Struct Biol*. 2013;23(5):660-668.
16. Kjaergaard CH, Qayyum MF, Wong SD, Xu F, Hemsworth GR, Walton DJ, Young NA, Davies GJ, Walton PH, Johansen KS, et al. Spectroscopic and computational insight into the activation of O_2 by the mononuclear Cu center in polysaccharide monooxygenases. *Proc Natl Acad Sci U S A*. June 2014.
17. Beeson WT, Phillips CM, Cate JHD, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc*. 2012;134(2):890-892.
18. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels*. 2013;6(1):41.
19. Vu VV, Beeson WT, Phillips CM, Cate JHD, Marletta MA. Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. *J Am Chem Soc*. 2014;136(2):562-565.
20. Isaksen T, Westereng B, Aachmann FL, Agger JW, Kracher D, Kittl R, Ludwig R, Haltrich D, Eijsink VGH, Horn SJ. A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. *J Biol Chem*. 2014;289(5):2632-2642.
21. Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JCN, Johansen KS, Krogh KBRM, Jørgensen CI, Tovborg M, Anthonsen A, et al. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci*. 2011;108(37):15079-15084.
22. Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunæs AC, Stenstrøm Y, Mackenzie A, Sørlie M, Horn SJ, Eijsink VGH. Cleavage of cellulose by a CBM33 protein. *Protein Sci*. 2011;20(9):1479-1483.
23. Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WGT, Ludwig R, Horn SJ, Eijsink VGH, Westereng B. Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. *Proc Natl Acad Sci*. 2014;111(17):6287-6292.
24. Vu VV, Beeson WT, Span EA, Farquhar ER, Marletta MA. A family of starch-active polysaccharide monooxygenases. *Proc Natl Acad Sci*. 2014;111(38):13822-13827.
25. Frommhagen M, Sforza S, Westphal AH, Visser J, Hinz SWA, Koetsier MJ, van Berkel WJH,

Gruppen H, Kabel MA. Discovery of the combined oxidative cleavage of plant xylan and cellulose by a new fungal polysaccharide monooxygenase. *Biotechnol Biofuels*. 2015;8(1):101.

26. Cannella D, Jørgensen H. Do new cellulolytic enzyme preparations affect the industrial strategies for high solids lignocellulosic ethanol production? *Biotechnol Bioeng*. 2014;111(1):59-68.
27. Corrêa TLR, dos Santos LV, Pereira GAG. AA9 and AA10: from enigmatic to essential enzymes. *Appl Microbiol Biotechnol*. 2016;100(1):9-16.
28. Sun FF, Hong J, Hu J, Saddler JN, Fang X, Zhang Z, Shen S. Accessory enzymes influence cellulase hydrolysis of the model substrate and the realistic lignocellulosic biomass. *Enzyme Microb Technol*. 2015;79-80:42-48.
29. Müller G, Várnai A, Johansen KS, Eijsink VGH, Horn SJ. Harnessing the potential of LPMO-containing cellulase cocktails poses new demands on processing conditions. *Biotechnol Biofuels*. 2015;8(1):187.
30. Hu J, Chandra R, Arantes V, Gourlay K, van Dyk JS, Saddler JN. The addition of accessory enzymes enhances the hydrolytic performance of cellulase enzymes at high solid loadings. *Bioresour Technol*. 2015;186:149-153.
31. Dimarogona M, Topakas E, Olsson L, Christakopoulos P. Lignin boosts the cellulase performance of a GH-61 enzyme from *Sporotrichum thermophile*. *Bioresour Technol*. 2012;110:480-487.
32. Koseki T, Mese Y, Fushinobu S, Masaki K, Fujii T, Ito K, Shiono Y, Murayama T, Iefuji H. Biochemical characterization of a glycoside hydrolase family 61 endoglucanase from *Aspergillus kawachii*. *Appl Microbiol Biotechnol*. 2008;77(6):1279-1285.
33. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. *Biotechnol Biofuels*. 2012;5(1):79.
34. Westereng B, Arntzen MØ, Aachmann FL, Várnai A, Eijsink VG, Agger JW. Simultaneous analysis of C1 and C4 oxidized oligosaccharides, the products of lytic polysaccharide monooxygenases (LPMOs) acting on cellulose. *J Chromatogr A*. 2016;Under review.
35. Beeson WT, Iavarone AT, Hausmann CD, Cate JHD, Marletta MA. Extracellular aldonolactonase from *Myceliophthora thermophila*. *Appl Environ Microbiol*. 2011;77(2):650-656.
36. Bennati-Granier C, Garajova S, Champion C, Grisel S, Haon M, Zhou S, Fanuel M, Ropartz D,

Rogniaux H, Gimbert I, et al. Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide monooxygenases secreted by *Podospora anserina*. *Biotechnol Biofuels*. 2015;8(1):90.

37. York WS, van Halbeek H, Darvill AG, Albersheim P. Structural analysis of xyloglucan oligosaccharides by ^{1}H-n.m.r. spectroscopy and fast-atom-bombardment mass spectrometry. *Carbohydr Res*. 1990;200:9-31.
38. Johansen KS. Discovery and industrial applications of lytic polysaccharide mono-oxygenases. *Biochem Soc Trans*. 2016;44(1):143-149.
39. Zhao Z, Liu H, Wang C, Xu JR. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics*. 2014;15(6):1-15.
40. Desmet T, Cantaert T, Gualfetti P, Nerinckx W, Gross L, Mitchinson C, Piens K. An investigation of the substrate specificity of the xyloglucanase Cel74A from *Hypocrea jecorina*. *FEBS J*. 2007;274(2):356-363.
41. Feng T, Yan KP, Mikkelsen MD, Meyer AS, Schols HA, Westereng B, Mikkelsen JD. Characterisation of a novel endo-xyloglucanase (XcXGHA) from *Xanthomonas* that accommodates a xylosyl-substituted glucose at subsite -1. *Appl Microbiol Biotechnol*. 2014;98(23):9667-9679.
42. Gilbert J, Fernando WGD. Epidemiology and biological control of *Gibberella zeae* / *Fusarium graminearum*. *Can J Plant Pathol*. 2004;26(4):464-472.
43. Garcia D, Barros G, Chulze S, Ramos AJ, Sanchis V, Marín S. Impact of cycling temperatures on *Fusarium verticillioides* and *Fusarium graminearum* growth and mycotoxins production in soybean. *J Sci Food Agric*. 2012;92(15):2952-2959.
44. Doohan FM, Brennan J, Cooke BM. Influence of climatic factors on *Fusarium* species pathogenic to cereals. *Eur J Plant Pathol*. 2003;109(7):755-768.
45. Park YB, Cosgrove DJ. Xyloglucan and its interactions with other components of the growing cell wall. *Plant Cell Physiol*. 2015;56(2):180-194.
46. Burton R a, Gidley MJ, Fincher GB. Heterogeneity in the chemistry, structure and function of plant cell walls. *Nat Chem Biol*. 2010;6(10):724-732.
47. Umemura M, Yuguchi Y. Conformational folding of xyloglucan side chains in aqueous solution from molecular dynamics simulation. *Carbohydr Res*. 2005;340(16):2520-2532.
48. Hori C, Gaskell J, Igarashi K, Samejima M, Hibbett D, Henrissat B, Cullen D. Genomewide analysis of polysaccharides degrading enzymes in 11 white- and brown-rot Polyporales provides insight into mechanisms of wood decay. *Mycologia*. 2013;105(6):1412-1427.

49. Loose JSM, Forsberg Z, Fraaije MW, Eijsink VGH, Vaaje-Kolstad G. A rapid quantitative activity assay shows that the *Vibrio cholerae* colonization factor GbpA is an active lytic polysaccharide monooxygenase. *FEBS Lett*. 2014;588(18):3435-3440.

50. Paspaliari DK, Loose JSM, Larsen MH, Vaaje-Kolstad G. *Listeria monocytogenes* has a functional chitinolytic system and an active lytic polysaccharide monooxygenase. *FEBS J*. 2015;282(5):921-936.

51. Garcia-Gonzalez E, Poppinga L, Fünfhaus A, Hertlein G, Hedtke K, Jakubowska A, Genersch E. *Paenibacillus larvae* chitin-degrading protein *Pl*CBP49 is a key virulence factor in American Foulbrood of honey bees. *PLoS Pathog*. 2014;10(7):e1004284.

52. Várnai A, Tang C, Bengtsson O, Atterton A, Mathiesen G, Eijsink VGH. Expression of endoglucanases in *Pichia pastoris* under control of the GAP promoter. *Microb Cell Fact*. 2014;13(1):57.

53. Wood TM. Preparation of crystalline, amorphous, and dyed cellulase substrates. In: *Methods in Enzymology*. Vol 160.; 1988:19-25.

54. Strohalm M, Kavan D, Novák P, Volný M, Havlícek V. mMass 3: a cross-platform software environment for precise analysis of mass spectrometric data. *Anal Chem*. 2010;82(11):4648-4651.

55. Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, Stanley P, Hart G, Darvill A, Kinoshita T, et al. Symbol nomenclature for graphical representations of glycans. *Glycobiology*. 2015;25(12):1323-1324.

FIGURES

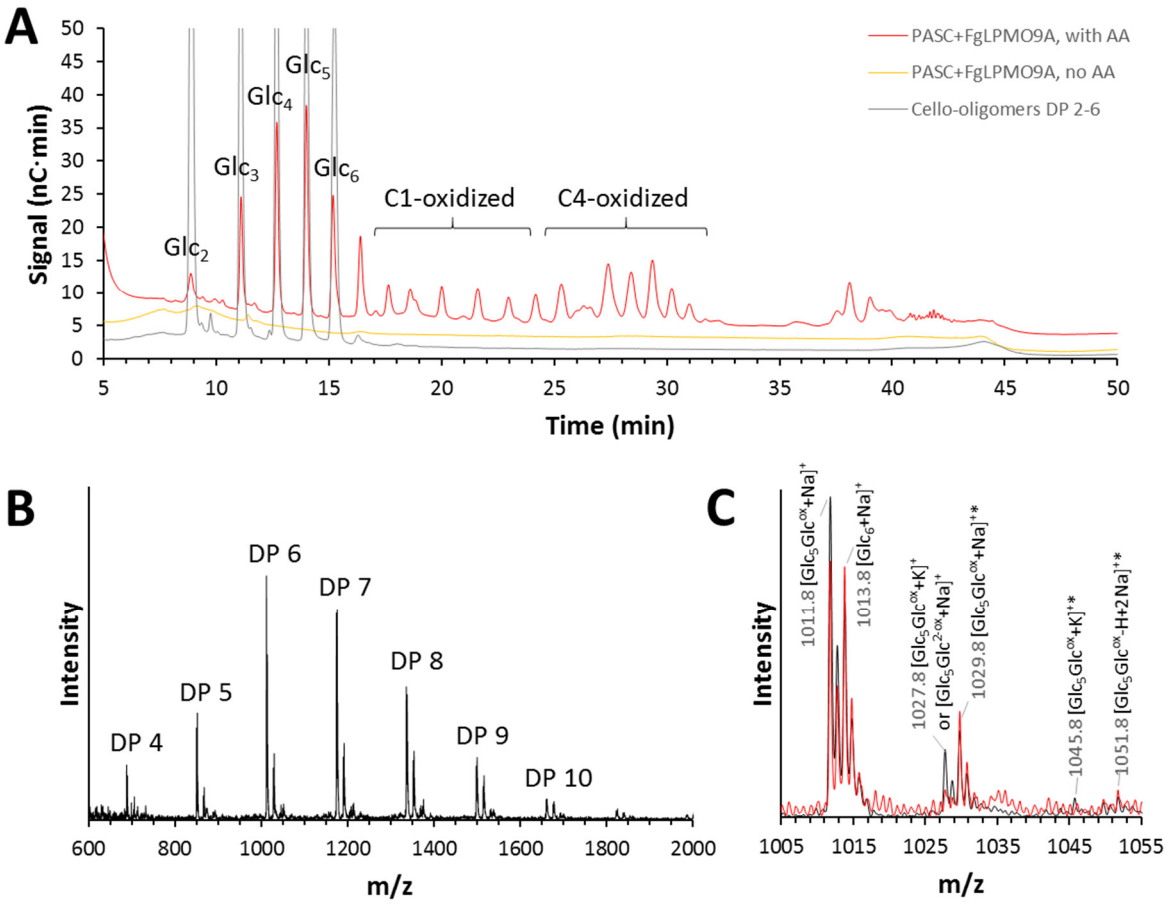


Figure 1. Reaction products generated by *Fg*LPMO9A from PASC. (A) HPAEC-PAD profiles of reaction mixtures containing *Fg*LPMO9A and PASC, with (red line) and without (yellow line) ascorbic acid; the grey line shows cello-oligosaccharide standards with DP 2-6. The produced oxidized and native cello-oligosaccharides are labeled in the figure and annotations are based on previous work [19,20]. (B) MALDI-ToF/MS spectrum of oligosaccharides released from PASC by *Fg*LPMO9A, ranging from DP 4-10. (C) Close-up of the DP 6 cluster before (black line) and after Na^+-saturation (red line). Intensity represents 1.5×10^4 arbitrary units for the highest peak in (B) and (C). Oxidation on either C1 or C4 is denoted with "ox"; oxidized products that are hydrated are marked with "*". Reactions were carried out in 20 mM Bis-Tris buffer pH 6.5 at 45 °C with 2 µM *Fg*LPMO9A, 4 mg/ml PASC and 1 mM ascorbate as external electron donor.

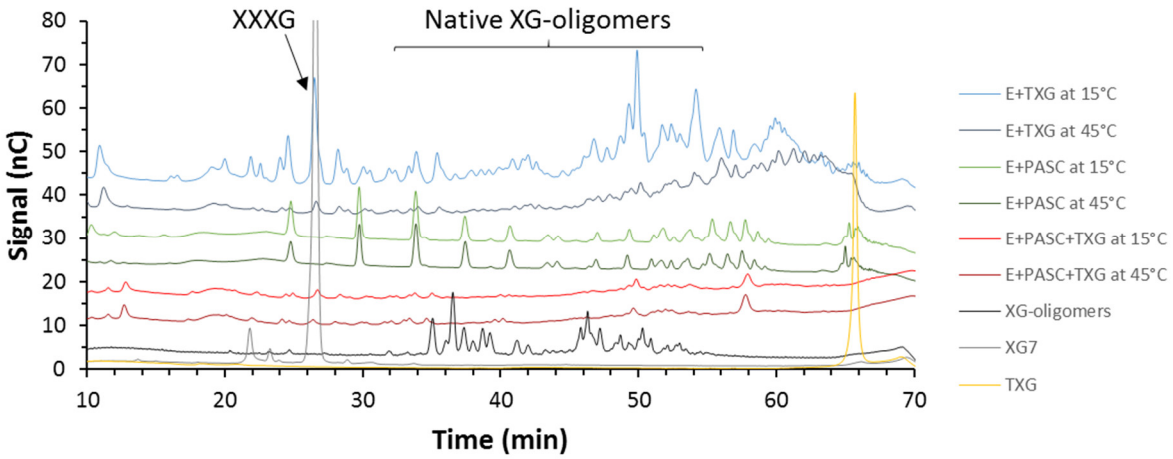


Figure 2. HPAEC-PAD chromatograms showing products obtained after *Fg*LPMO9A action on tamarind xyloglucan (TXG) at 15 °C (light blue line) or 45 °C (dark blue line), PASC at 15 °C (light green line) or 45 °C (dark green line), PASC coated with TXG at 15 °C (light red line) or 45 °C (dark red line), as well as a xyloglucan-oligosaccharide standard (black line), a xyloglucan heptamer (XXXG) standard (grey line) and TXG (yellow line). Control reactions showed that neither native nor oxidized xyloglucan oligosaccharides were generated by *Fg*LPMO9A in the absence of an electron donor. Reactions were carried out in 20 mM Bis-Tris buffer pH 6.5 with 2 µM *Fg*LPMO9A, 1 mg/ml TXG or 4 mg/ml PASC or mix of 1 mg/ml TXG and 4 mg/ml PASC and 1 mM ascorbate as external electron donor.

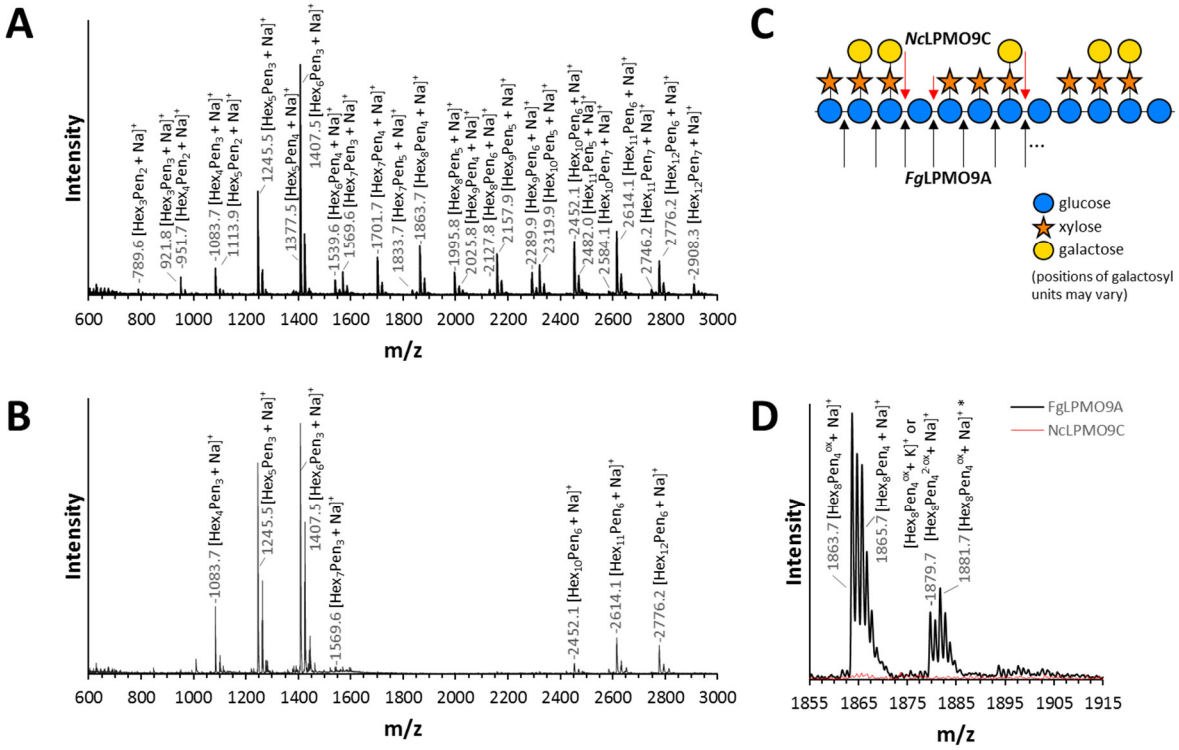


Figure 3. Reaction products generated from tamarind xyloglucan (TXG). (A) MALDI-ToF/MS spectrum showing product profile from TXG by *Fg*LPMO9A with oxidized product peak annotations. (B) MALDI-ToF/MS spectrum showing product profile from TXG by *Nc*LPMO9C with oxidized product peak annotations. Note the much more complex product profile obtained with *Fg*LPMO9A. (C) Illustration of the structure of a fragment of TXG (blue circle, glucose; orange star, xylose; yellow circle, galactose [55]) and possible cleavage sites for *Fg*LPMO9A (black arrows, and ellipsis signify continuity) and *Nc*LPMO9C (red arrows). (D) Close-up of the Hex_8Pen_4 cluster showing the sodium or potassium adduct of native and oxidized products; hydrated oxidized products are labeled with "*". Intensity represents 5.9×10^4 arbitrary units (a.u.) for the highest peak in (A), 9.1×10^4 a.u. in (B) and 1.2×10^4 a.u. in (D). Reactions were carried out in 20 mM Bis-Tris buffer pH 6.5 at 15°C with 1 mg/ml TXG, 2 µM *Fg*LPMO9A or *Nc*LPMO9C and 1 mM ascorbate as external electron donor. Abbreviations: Hex, hexose (+ 162 Da); Pen, pentaose (+ 132 Da); ox, oxidized.

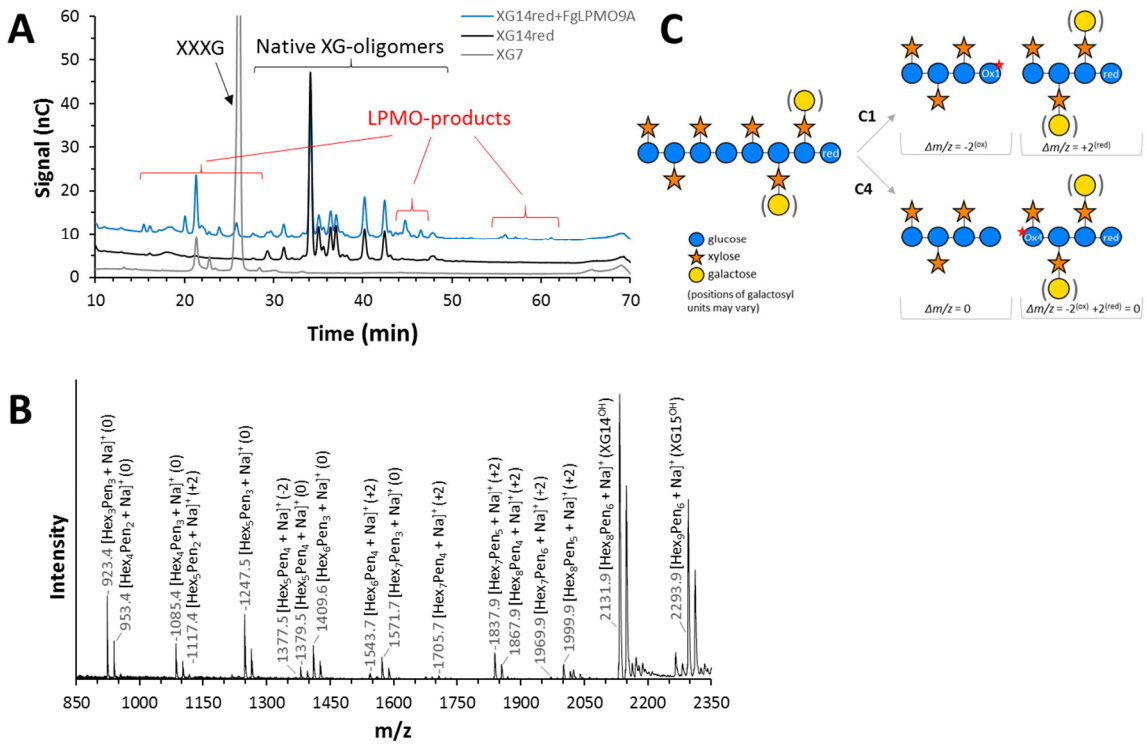


Figure 4. Analysis of products generated from XG14^{OH}. (A) HPAEC-PAD chromatograms showing products generated upon incubation of XG14^{OH} with *Fg*LPMO9A (blue line), as well as non-treated XG14^{OH} (black line) and the xyloglucan heptamer (XXXG) standard (grey line). (B) MALDI-ToF/MS spectrum showing products generated from XG14^{OH} by *Fg*LPMO9A with C1- and C4-oxidized product peak annotations. Numbers in parenthesis show Δ*m/z* compared to the mass of the native XG-oligosaccharide products. Intensity represents 9 × 10^4 arbitrary units for the highest peak. Reactions without ascorbic acid showed a peak of XG14^{OH} and several minor species containing one or more L, whereas no smaller products were present. (C) The substrate, XG14^{OH} (blue circle, glucose; orange star, xylose; yellow circle, galactose [55]), and comparison of the C1- and C4-oxidation products and their masses to native ones. Parenthesis surrounding galactosyl-units signify that the position of these units may vary. Reactions were carried out in 20 mM Bis-Tris buffer pH 6.5 at 15 °C with 1 mg/ml XG14^{OH}, 2 µM *Fg*LPMO9A and 1 mM ascorbate as external electron donor. Abbreviations: Hex, hexose (+ 162 Da); Pen, pentaose (+ 132 Da); "red", the position of reduction; "Ox1" and "Ox4", position of C1- and C4-oxidation.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site.

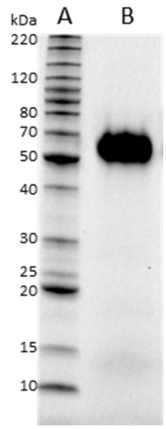


Figure S1. SDS-PAGE of purified *Fg*LPMO9A. Lanes: (A) Benchmark standard protein molecular weight marker (MW, values labelled), (B) purified *Fg*LPMO9A fraction after size exclusion chromatography.

Concluding remarks and perspectives

The work presented in this thesis includes the investigation of the secretome of *Aspergillus nidulans* growing on various starches as well as expansion of knowledge on novel fungal lytic polysaccharide monooxygenases, their activities and substrate binding.

To design better and more efficient enzyme cocktails for plant biomass degradation, a good starting point is to study the transcriptomes and secretomes of microorganisms grown on desired biomass. While the main focus is on lignocellulosic material, much fewer studies involve fungal degradation of starch. The filamentous fungus *A. nidulans* was shown to secrete enzymes with different profiles depending on the type of starch and its resistance to hydrolytic degradation (Paper I). It also shows that novel starch active LPMOs are amongst the most abundant CAZymes secreted together with a range of other oxidative enzymes, which highlights the *in vivo* role of oxidative enzyme in starch degradation by fungi and the regulation of amylolytic enzyme activities, e.g. at what time of growth they are secreted, co-secretion with other enzymes, composition of enzymes, etc. We have also identified AmyR, starch degradation transcriptional regulators, upstream the two LPMOs, providing the first evidence for a co-regulation of GHs and LPMOs targeting a complex polysaccharide. This study also shows that a variety of cell wall degrading enzymes are co-secreted with amylolytic enzymes. A possible rationale for this is the caging of some of the starch in cell-wall matrices, which is overcome by co-secretion of cell wall degrading enzymes. This insight promotes the design of better enzyme cocktails for more exhaustive and efficient degradation of starch, especially from more crude sources, where the addition of cell-wall degrading enzymes may enhance the feasibility of the process. Undoubtedly, studies on industrially realistic biomass substrates are required.

The presence of carbohydrate binding modules (CBMs) in LPMOs is an important feature that needs to be taken into account when considering engineering LPMOs to improve their catalytic function. An understanding of the extent of CBM-mediated LPMO binding to substrates and the impact of that on activity is an interesting avenue to study. Currently, the information on the role of CBMs in the activity of LPMOs is scarce and it is evident that more detailed analyses of CBMs influence are needed. Paper II gives first insight into starch-specific LPMOs binding to starch and starch related substrate. The efficiency of some LPMOs may be boosted by the engineering of CBMs that may enhance their substrate binding properties.

The identification of new LPMOs and their families with new enzyme activities is significantly expanding knowledge of biomass degradation in nature and its application in industrial processes.

A novel AA9 LPMO from the filamentous fungus *Fusarium graminearum* was characterized in Paper III. For the first time, LPMO was showed to cleave xyloglucan at both substituted and unsubstituted glucosyl units and was also active on cellulose. Further studies to understand whether and how LPMOs synergize with each other and also with other enzymes could further improve biomass decomposition efficiency. It is possible to envisage tailoring the composition of hydrolytic cocktails with multiple LPMOs

Since the discovery of LPMOs, the research on these enzymes has progressed but still a lot of questions are remaining: about their chemistry at the copper active site, mechanism, better solutions for electron sources, utilization in the biorefineries. Moreover, a robust kinetic assay would help to engineer LPMOs that have higher catalytic activity.

CONFERENCE CONTRIBUTIONS

Poster presentations in conferences:

- **Nekiunaite** L, Vaaje-Kolstad G, Svensson B, Abou Hachem, M. Exploring the secretomes of starch and chitin degrading fungi. Gordon Research Conference (GRC) on Cellulosomes, Cellulases & Other Carbohydrate Modifying Enzymes, Proctor Academy in Andover, New Hampshire, USA. August 2013.
- **Nekiunaite** L, Vaaje-Kolstad G, Svensson B, Abou Hachem, M. Exploring the secretomes of starch degrading fungi. Fifth Symposium on the Alpha-Amylase Family (ALAMY_5), Smolenice, Slovakia. October 2013.
- **Nekiunaite** L, Vaaje-Kolstad G, Svensson B, Abou Hachem, M. Exploring the secretomes of starch degrading fungi. 10th Workshop in Protein.DTU, Kgs. Lyngby, Denmark. November 2013.
- **Nekiunaite** L, Vaaje-Kolstad G, Svensson B, Øverlie Arntzen M, Abou Hachem, M. Exploring the secretomes of starch degrading fungi. 11th Carbohydrate Bioengineering Meeting (CBM11), Espoo, Finland. May 2015. *Second prize for the poster.*